

High Efficiency Gene Correction in Hematopoietic Cells by Donor Template-free CRISPR/Cas9 Genome Editing

Dissertation

Vom Fachbereich Biologie
der Technischen Universität Darmstadt

zur Erlangung des akademischen Grades
eines Doctor rerum naturalium

genehmigte Dissertation von

M.Sc. Duran Sürün
aus Pazarcik

1. Referentin: Prof. Dr. Beatrix Süß
2. Referentin: Prof. Dr. M. Cristina Cardoso
Externer Referent: Prof. Dr. Harald von Melchner

Tag der Einreichung: 01.11.2017

Tag der mündlichen Prüfung:
06.02.2018

Darmstadt 2018

D17

Teile der vorliegenden Arbeit wurden in der folgenden Publikation veröffentlicht:

Sürün D, Schwäble J, Tomasovic A, Roy E, Stein S, Kurrle A, Kühn A, von Melchner H and Schnütgen F

High Efficiency Gene Correction in Hematopoietic Cells by Donor-Template-free CRISPR/Cas9 Genome Editing

Molecular Therapy - Nucleic Acids, in Revision

Table of contents

Abbreviations and Definitions	VII
Summary	1
Zusammenfassung	4
1. Introduction	7
1.1 The hematopoietic system.....	7
1.2 Primary immunodeficiency diseases	9
1.3 Molecular background of PIDs	11
1.4 Hematopoietic stem cell transplantation and gene therapy	12
1.5 Gene transfer vector systems.....	14
1.5.1 Non-viral vectors.....	14
1.5.2 Viral vectors	16
1.5.3 Non-integrating viral vectors	17
1.5.4 Integrating viral vectors.....	18
1.5.5 Vectors of choice for gene replacement therapy in hematopoietic stem cells	19
1.6 Clinical gene therapy trials for the treatment of PIDs	21
1.7 Chronic granulomatous disease	23
1.8 Gene therapy for X-CGD	24
1.9 Site-specific endonucleases	26
1.9.1 Meganucleases.....	26
1.9.2 Zinc-finger nucleases.....	27
1.9.3 Transcription activator-like effectors nucleases	27
1.9.4 Clustered regularly interspaced short palindromic repeats/Cas9.....	28
1.10 Molecular outcomes of genome editing	29
1.11 Aim of this work	30
2. Materials and Methods.....	31
2.1 Material.....	31

2.1.1	Software and online tools	31
2.1.2	Hardware and consumables	31
2.1.3	Plastic ware and consumables	32
2.1.4	Chemicals and reagents	33
2.1.5	Buffers and solutions	35
2.1.6	Enzymes.....	36
2.1.7	Commercially available kits	36
2.1.8	FACS antibodies.....	36
2.1.9	Western blot antibodies	37
2.1.10	Cytokines.....	37
2.1.11	Plasmids	37
2.1.12	Cell lines and culture media.....	40
2.1.13	Bacterial stains	40
2.1.14	Oligonucleotides	41
2.2	Methods of molecular biology.....	43
2.2.1	Cultivation of bacteria	43
2.2.2	Cloning processes	43
2.2.3	Transformation of competent bacteria	46
2.2.4	Plasmid preparation.....	47
2.2.5	Agarose gel electrophoresis	47
2.2.6	Gel extraction of DNA fragments	48
2.2.7	Polymerase chain reaction (PCR).....	48
2.2.8	PCR product extraction.....	49
2.2.9	Nucleic acid sequencing	49
2.2.10	Isolation of genomic DNA from cells	49
2.2.11	Surveyor Assay	50
2.2.12	Southern Blot.....	51
2.2.13	Western Blot	52

2.3	Cell culture and virological methods	54
2.3.1	Cultivation of cell lines	54
2.3.2	Cell counting and determination of cell viability	54
2.3.3	Make WEHI-3B conditioned media	55
2.3.4	Freezing and thawing of cultured cell	55
2.3.5	Nucleofection	55
2.3.6	Lentiviral vector production	56
2.3.7	Titration of vector particles	57
2.3.8	Flow cytometry and cell sorting	58
2.3.9	Dihydrorhodamine 123 (DHR) reduction assay	58
2.3.10	Statistical analysis	59
3.	Results	60
3.1	Genome editing by homology-directed repair (HDR) versus non-homologous end joining (NHEJ)	60
3.2	Reconstitution of a point mutated <i>Hprt</i> gene by genome editing in murine hematopoietic (32D) cells	61
3.3	NHEJ-mediated reconstitution of a mutated <i>EGFP</i> reporter gene in 32D cells	63
3.4	NHEJ mediated reconstitution of a mutated EGFP reporter transgene in PLB-985 human promyelocytic leukemia cells	67
3.5	Reconstitution of point mutated <i>CYBB</i> gene expressed in X-CGD-PLB leukemia cells	73
3.6	Estimating on target mutation efficiency at the endogenous CYBB locus in PLB cells	84
4.	Discussion	87
5.	Conclusions and Outlook	96
1.	References	98
	Acknowledgments	107
	Curriculum vitae	108

Publications and Abstracts	109
Ehrenwörtliche Erklärung:.....	111

List of figures

Figure 1: The hematopoietic system.....	8
Figure 2: Mutation type of the selected examples of PIDs.....	12
Figure 3: In vivo and ex vivo gene therapy concepts.....	13
Figure 4: Overview of vectors used in gene therapy clinical trials	14
Figure 5: Various nonviral gene delivery systems (modified from Manjila et al. 2013)	15
Figure 6: Transposon vector system for stable gene delivery (modified from Ivics et al. 2011 ³⁷).....	16
Figure 7: Analysis of integration frequencies around transcriptional start sites (TSS)	19
Figure 8: Mechanisms of the insertional mutagenesis mediated by the retroviral vectors	20
Figure 9: NADPH oxidase complexes in phagocytic cells	24
Figure 10: TALEN-assisted <i>Hprt</i> gene editing	61
Figure 11: Analysis of INDEL distribution	62
Figure 12: TALEN-assisted random <i>Hprt</i> gene repair.....	63
Figure 13: <i>EGFP</i> mutagenesis in 32D-GFP ⁺ cells	64
Figure 14: Indels induced by NHEJ in 32D-GFP negative cells enriched by FACS and subcloned by limiting dilution	65
Figure 15: <i>EGFP</i> repair frequency in selected 32DG subclones	66
Figure 16: Generation of PLB-985 cells expressing dual color reporter transgenes.	68
Figure 17: Single copy integration of SBmGW lentiviruses in PLB cells.....	69
Figure 18: EGFP repair efficiency in PLB-985 cells expressing dual color reporters	70
Figure 19: mEGFP alleles in IDLV infected SBmGW-PLB cells	72
Figure 20: FACS profiles of XCGD cells expressing wild-type and mutant <i>CYBB</i> cDNAs	74
Figure 21: CYBB protein expression in XCGD-SBwtCW and SBmCW cells detected by Western blot.....	75
Figure 22 Repair of CYBB mutations in XCGD-PLB cells	76
Figure 23: FACS profile of sorted CYBB ⁺ XCGD cells.....	77
Figure 24: Indel analyses of the shot-gun clones and single cells.....	79
Figure 25: FACS analyses of the single cell clones form SB45CW, SB54CW and SB173CW	80

Figure 26: Differentiation of XCGD cells.....	81
Figure 27: NADPH oxidase activity in CYBB corrected XCGD cells.....	82
Figure 28: Functional analysis of the SB173CW-C24 clones	83
Figure 29: On target mutation efficiency at the endogenous exon 3 and exon 6 CYBB gene	85
Figure 30: Indel analyses of the shot-gun clones from endogenous CYBB locus.....	86

List of tables

Table 1: Selected examples of PIDs and their prevalence ^{13,14}	10
Table 2: Viral vectors and their main properties	17
Table 3: Clinical trials of HSC-based gene-therapy in PIDs.....	22
Table 4: Summary of gene therapy trials for X-CGD including myelosuppressive strategies	25
Table 5: Examples of applications of genome editing to therapeutic model disease.	29
Table 6: Selected <i>CYBB</i> disease mutations and nucleotide sequences of the <i>CYBB</i> -targeting sgRNAs	73
Table 7: Types of indels recovered from sorted <i>CYBB</i> ⁺ XCGD cells	78
Table 8: Computational Transmembrane sequence analysis from XCGD-SB173CW single cell clones	81
Table 9: Theoretical frequency of gene repair	94

Abbreviations and Definitions

AAV	adeno-associated viruses
ADA-SCID	adenosine deaminase-deficient severe combined immunodeficiency
Amp	ampicillin
APC	allophycocyanin
bp	base pairs
BSA	bovine serum albumin
°C	degree Celsius
CGD	chronic granulomatous disease
CLP	common lymphoid progenitors.
CMP	common myeloid progenitors.
CRISPR	Clustered regularly interspaced short palindromic repeats
CYBB	Cytochrome b-245 heavy chain
DNA	Deoxyribonucleic acid
DSB	DNA double strand break
DHR	dihydrorhodamine123
DMEM	Dulbecco`s modified Eagle medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dsDNA	double-stranded deoxyribonucleic acid
<i>E. coli</i>	Escherichia coli
EDTA	ethylenediaminetetraacetic acid
EGFP	enhanced green fluorescent protein
Env	envelope protein
et al.	and others
FACS	fluorescence activated cell sorting
FCS	fetal calf serum
FITC	fluorescence isothiocyanate
FSC	forward scatter
g	gram
Gag	group-specific-antigen
GOI	gene of interest
GMP	granulocyte-macrophage progenitors
GvHD	graft-versus-host disease
h	hour
HAT medium	hypoxanthine-aminopterin-thymidine medium
H ₂ O ₂	hydrogen peroxide
HIV-1	Human immunodeficiency virus-1
HDR	homology directed repair

HLA	human leukocyte antigens
HOCl	hypochlorite ion
HPRT	hypoxanthine-guanine phosphoribosyltransferase
HSC	hematopoietic stem cells
HSCT	hematopoietic stem cell transplantation
HSPC	hematopoietic stem and progenitor cells
IDLV	Integrase-defective lentiviral
IFN	interferon
iL2RG	interleukin-2 receptor gamma chain
IN	integrase
IRES	internal ribosomal entry site
kb	kilobase pair
KI	<i>knock-in</i>
KO	<i>knock-out</i>
l	liter
LB	Luria Broth
LTR	long terminal repeat
M	molar
MA	matrix
MEP	megakaryocyte-erythrocyte progenitor
mEGFP	mutationally inactivated EGFP
MLV	murine leukaemia virus
MOI	multiplicity of infection
MPP	multipotent progenitors
mRNA	messenger RNA
NADPH	nicotinamide adenine dinucleotide phosphate
NEB	New England Biolabs
NHEJ	nonhomologous end joining
NK cell	natural killer cell
NMD	nonsense mediated decay
nt	nucleotide
O ₂ ⁻	superoxide anion
OH [·]	hydroxyl radical
ORF	open reading frame
PB	PiggyBac
PBS	primer binding site or phosphate buffered saline
PCR	polymerase chain reaction
PE	R-Phycoerythrin
PEI	Polyethylenimin
phox	phagocytic oxidase
PID	primary immunodeficiency diseases

PMA	phorbol 12-myristate 13-acetate
Pol	polymerase
psi	packaging signal of retroviral genomic RNA
RGN	RNA-guided nucleases
RNA	ribonucleic acid
RNase	ribonuclease
ROS	reactive oxygen species
rpm	revolutions per minute
RPMI	Roswell Park Memorial Institute - culture medium
RT	room temperature or reverse transcriptase
SA	splice acceptor
SB	Sleeping Beauty
SD	slice donor
SDS	sodium dodecyl sulfate
sec	second
SFFV	spleen focus-forming virus
SIN	self-inactivating
SSC	side scatter
SSDNA	single-stranded Deoxyribonucleic acid
SV40	Simian virus 40
TagBFP	blue fluorescent protein
TALEN	transcription activator–like effector nucleases
TIR	terminal inverted repeats
TU	transducing units or Technical University
V	volt
VSV-G	glycoprotein of vesicular stomatitis virus
v/v	Volume/volume
WAS	Wiskott-Aldrich syndrome
WASp	Wiskott-Aldrich syndrome protein
w/v	weight/volume (Volume concentration)
WPRE	Woodchuck hepatitis virus posttranscriptional regulatory element
WT	wild type
X-CGD	X-linked chronic granulomatous disease
X-SCID	X-linked severe combined immunodeficiency
ZFN	Zinc-finger nuclease
ZFP	zinc-finger protein
μ	micro-
6-TG	6-thioguanine

Summary

A significant fraction of inherited monogenic disorders are caused by patient-specific mutations dispersed over the entire locus of the affected gene. Although correcting these mutations by introducing healthy gene copies into the genome of the diseased cells proved effective in several clinical gene therapy trials and with more advanced vectors safety and efficacy could be improved, insertional mutagenesis and unregulated expression of genes deprived of their endogenous control elements remains a concern when using randomly integrating vectors. As has been shown repeatedly in clinical trials random vector insertions are susceptible to epigenetic silencing and can cause cancer by the activation of adjacent proto-oncogenes.

The development of genome editing tools capable of modifying any prespecified genomic sequence with unprecedented accuracy opened up a wide range of new possibilities in gene manipulation including targeted gene repair. In particular, CRISPR/Cas9 system, a prokaryotic adaptive immune system and its swift repurposing for genome editing was widely adopted as the hitherto simplest genome editing tool. In combination with a single guide RNA (sgRNA) the Cas9 endonuclease generates DNA double strand breaks (DSBs) at prespecified genomic loci that are repaired either by homology directed repair (HDR) or nonhomologous end joining (NHEJ).

Correction of human disease mutations by this technology has been thus far largely based on homologous recombination requiring an exogenous donor template along with RNA guided (gRNA) Cas9 endonucleases (RGNs). In most applications, RGNs and templates were delivered to the diseased cells by electroporation of several plasmids each expressing one of the functional components needed for targeted gene modification. However, transducing the functional components required for homology directed repair (HDR) on different plasmids and considering that electroporation is quite harmful to the target cells, only a small fraction of the cells survive transfection and even fewer retain all functional components. As a result, the number of gene corrected cells is usually quite low and reduced even further by the inherent bias of the cell's double strand break (DSB) machinery towards NHEJ.

This thesis explores the efficiency of gene repair by NHEJ in hematopoietic cells harboring patient specific point mutations in the Cytochrome b-245 heavy chain gene (*CYBB*) whose inactivation causes chronic granulomatous disease (X-CGD), - a life-threatening immunodeficiency disorder. Although in contrast to HDR, NHEJ is error prone, the present work was based on the theoretical assumption that about, one-third of the insertions/deletions (indels) associated with NHEJ should restore the open reading frame (ORF) disrupted by a particular disease mutation. This would lead to a significant number of ORF reconstitutions of which some, depending on the position and type of the original mutation, should either completely or partially recover protein function. Moreover, donor template free delivery of RGNs on one rather than multiple expression vectors by lentiviral infection was expected to improve gene repair efficiencies and to reduce toxicity of gene transduction.

In initial experiments designed to determine the efficiency of gene repair by NHEJ 32D hematopoietic cells expressing four different EGFP reporter transgenes harboring N-terminal frameshift mutations were nucleofected each with Cas9 and corresponding sgRNAs. Consistent with previous genome editing protocols involving transfection, gene repair efficiency was low, ranging from 2.3% to 5.5%.

Similar testing was performed in human PLB-985 leukemia cells expressing one copy of a mutationally inactivated EGFP reporter (mEGFP). However, to increase transduction rates and ensure transient RGN expression, the RGNs were delivered by integration defective lentiviruses (IDLVs). Unlike transfection IDLV delivery of RGNs yielded high on-target mutation rates leading to mEGFP repair rates of up to 27%. Collectively, the results demonstrate that mEGFP repair efficiency improved by one order of magnitude after changing the RGN delivery protocol from plasmid nucleofection to IDLV infection.

This strategy was tested further in PLB cells harboring bona fide disease mutations. For this, four X-CGD-patient specific *CYBB* mutations including two frameshift, one nonsense and one missense mutation were individually transduced into *CYBB* null PLB cells (XCGD-PLB). While subsequent delivery of the corresponding RGNs effectively repaired the frameshift mutations in up to 10% of the treated cells, the repair

efficiency of the nonsense and missense mutations was with less than 2% rather ineffective.

As about 20 - 25% of most inherited blood disorders are caused by frameshift mutations, the results of this thesis suggest that up to a quarter of all patients suffering from monogenic blood disorders could benefit from a gene therapy employing personalized, donor-template free RGNs.

Zusammenfassung

Ein signifikanter Anteil hereditärer, monogenetischer Erkrankungen wird durch patientenspezifische Mutationen verursacht, die über den gesamten Locus des betroffenen Gens verteilt sind. Obwohl sich in mehreren klinischen Studien die Korrektur dieser Mutationen durch die Einführung gesunder Genkopien in das Genom der erkrankten Zellen bewährt hat und mit fortschrittlicheren Vektoren die Sicherheit und Wirksamkeit verbessert werden kann, bleibt die Insertionsmutagenese und die unregulierte Expression von Genen außerhalb des Einflusses ihrer endogenen Kontrollelemente ein Problem bei der Verwendung von zufällig integrierenden Vektoren. Wie in klinischen Studien wiederholt gezeigt wurde, sind diese Vektoren anfällig für epigenetisches Silencing (Gen-Stillegung) und können durch Aktivierung benachbarter Protoonkogene Krebs verursachen.

Die Entwicklung von Genom-Editierungs-Technologien, die in der Lage sind, jede vorher festgelegte genomische Sequenz mit bisher unerreichter Präzision zu modifizieren, eröffnete eine breite Palette neuer Möglichkeiten in der Genmanipulation einschließlich gezielter Genreparatur. Insbesondere das prokaryotische adaptive Immunsystem CRISPR/Cas9, fand nach seiner eleganten Umfunktionierung zur Editierung doppelsträngiger DNA wurde als das bisher einfachste Genom-Editierwerkzeug breite Akzeptanz in der wissenschaftlichen Gemeinschaft.

In Kombination mit einer einzigen guide RNA (sgRNA) erzeugt die Cas9-Endonuklease DNA-Doppelstrangbrüche (DSBs) an vordefinierten genomischen Loci. Diese DSB können entweder durch Homologie-gerichtete Reparatur (HDR) oder nicht-homologe Verbindung der Strangenden (NHEJ) repariert werden.

Die Korrektur menschlicher Krankheitsmutationen durch diese Technologie basierte bislang weitgehend auf Homologie-gerichtete Reparatur, die neben RNA-geführten (gRNA) Cas9-Endonukleasen (RGNs) eine exogene homologe Donor-Matrize erfordert. In den meisten Anwendungen wurden RGNs und Matrizen durch Elektroporation in die erkrankten Zellen eingebracht, die auf mehreren Plasmiden kodiert sind, welche jeweils eine der funktionellen Komponenten exprimiert, die für eine gezielte Genmodifikation benötigt werden. Aufgrund dessen, dass die für die Homologie-gerichtete Reparatur (HDR) erforderlichen Komponenten auf

verschiedenen Plasmiden liegen und die Elektroporation für die Zielzellen ziemlich schädlich ist, überlebt nur ein kleiner Teil der Zellen die Transfektion und noch weniger beinhalten alle funktionellen Komponenten für die HDR. Infolgedessen ist die Anzahl genkorrigierter Zellen in der Regel recht niedrig und wird durch die inhärente Bevorzugung von NHEJ seitens der Doppelstrangbruch-Maschinerie in Stammzellen noch weiter reduziert.

In dieser Arbeit wird die Effizienz der Genreparatur durch NHEJ in hämatopoetischen Zellen mit patientenspezifischen Punktmutationen im Cytochrom b-245-Gen (CYBB) untersucht, dessen Inaktivierung die lebensbedrohliche Immunschwächekrankheit chronische Granulomatose (X-CGD) verursacht. Obwohl die NHEJ im Gegensatz zu HDR fehleranfällig ist, basierte die vorliegende Arbeit auf der theoretischen Annahme, dass ungefähr ein Drittel der mit NHEJ assoziierten Insertionen/Deletionen (Indels) den korrekten Leserahmen (ORF) wiederherstellen werden, der initial durch eine bestimmte Krankheitsmutation verschoben war. Dies würde zu einer signifikanten Anzahl von ORF-Rekonstitutionen führen, von denen einige, abhängig von der Position und Art der ursprünglichen Mutation, entweder die Proteinfunktion vollständig oder teilweise wiederherstellen sollten. Darüber hinaus wurde erwartet, dass die Donor-Matrize-freie Abgabe von RGNs auf einem statt von mehreren Expressionsvektoren durch lentivirale Infektion die Effizienz der Genreparatur verbessern und die Toxizität der Gentransduktion verringert.

Beim ersten Experimenten zur Bestimmung der Effizienz der Genreparatur durch NHEJ wurden die 32D hämatopoetische Zellen, die vier verschiedene EGFP-Reportertransgene mit N-terminalen Frameshift-Mutationen exprimierten, jeweils mit Cas9 und entsprechenden sgRNAs nukleofektiert. In Übereinstimmung mit früheren Genom-Editierprotokollen, die eine Transfektion beinhalten, war die Effizienz der Genreparatur gering und lag zwischen 2,3% und 5,5%.

Ähnliche Tests wurden in humanen PLB-985 Leukämiezellen durchgeführt, die eine Kopie eines mutationsinaktivierten EGFP-Reporters (mEGFP) exprimierten. Um die Transduktionsraten zu erhöhen und transiente RGN-Expression zu gewährleisten, wurden die RGNs durch Integrations-defekte Lentiviren (IDLVs) transduziert. Im Gegensatz zur Nucleofektion ergab die IDLV-Infektion von RGNs eine hohe On-

Target-Mutationsraten, was zu mEGFP-Reparaturraten von bis zu 27% führte. Insgesamt zeigen die Ergebnisse, dass sich die mEGFP-Reparatureffizienz um eine Größenordnung verbessert hat, nachdem das RGN Transduktionsprotokoll von Plasmid-Nukleofektion zu IDLV-Infektion geändert wurde.

Anschließend wurde die Strategie bei der hereditären septischen Granulomatose (X-CGD) die durch Mutationen im Cytochrome b-245 beta polypeptide (CYBB) Gen entstehen, in PLB985 Zellen getestet. Dazu wurden vier X-CGD-Patientenspezifische CYBB-Mutationen mit zwei Frameshift-, einer Nonsense- und einer Missense-Mutation einzeln in CYBB-Null-PLB-Zellen (XCGD-PLB) eingebracht. Während der nachfolgenden Transduktion der Zellen mit den entsprechenden IDLV-RGNs, wurden effektiv bis zu 10% die Frameshift-Mutationen korrigiert. Die Reparatureffizienz bei den Nonsense- und Missense-Mutationen war mit weniger als 2% jedoch eher ineffektiv.

Da etwa 20 – 25% der meisten vererbten Bluterkrankungen durch Frameshift-Mutationen verursacht werden, legen die Ergebnisse dieser Arbeit nahe, dass ein Viertel aller an monogenen Blutkrankheiten leidenden Patienten von einer Gentherapie profitieren könnten, die personalisierte, Donor-Template-freie RGNs verwendet.

1. Introduction

1.1 The hematopoietic system

The first theory of hematopoiesis was postulated by A. Maximow in 1909, stating that hematopoiesis accounts for the lifelong of formation of blood cells and plasma from hematopoietic stem cells (HSC)¹. Healthy adults produce billions of blood cells each day to replace an equal number of senescent or apoptotic cells removed from circulation^{2,3}. The HSC are ultimately responsible for blood cell renewal during which they differentiate via a series of precursor stages into terminally differentiated cells that acquire specific functions⁴.

Based on the ability to reconstitute hematopoiesis of irradiated recipients, HSC can be divided into long-term (LT-HSCs) and short-term hematopoietic stem cells (ST-HSCs). LT-HSCs possess unlimited self-renewal capacity and can therefore maintain hematopoiesis for an entire lifespan. ST-HSCs have limited self-renewal capacity and reconstitute hematopoiesis only transiently^{4,5}. Both cell types have their biological niche within the bone marrow⁶. As a progeny of LT-HSCs, ST-HSCs differentiate into multipotent progenitors (MPPs) which in turn give rise to the common myeloid (CMP) and lymphoid (CLP) progenitors. CMPs that differentiate into granulocyte-macrophage progenitors (GMPs) give rise to granulocytes, monocytes, dendritic cells, whereas CMPs differentiating into megakaryocyte-erythrocyte progenitor (MEPs) give rise to erythrocytes and platelets. CLPs generate B- and T-cells lymphocytes as well as NK-cells(see Figure 1)^{3,4,6}.

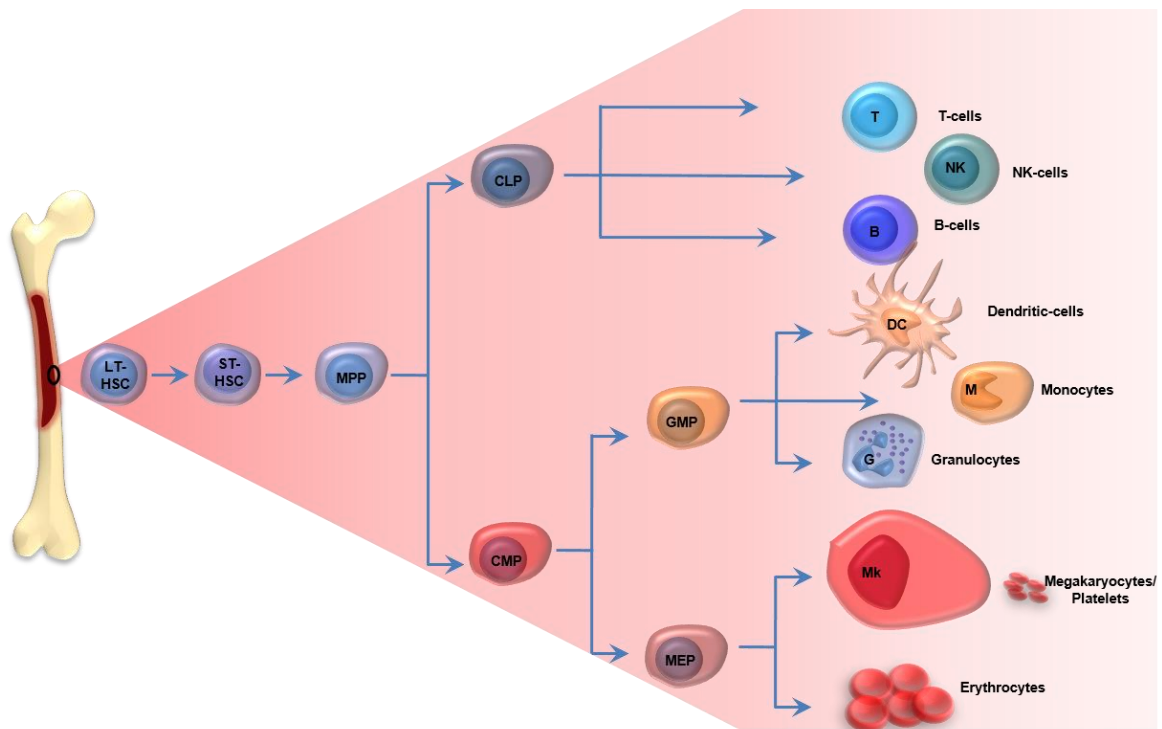


Figure 1: The hematopoietic system

During differentiation, HSCs increasingly lose their potential for self-renewal and their proliferation rate increases strongly, thus resulting in the enormous expansion capacity of blood cells.

For example, erythrocytes deliver oxygen to tissues and organs, platelets assist blood clotting during wound healing and tissue repair, and leukocytes are components of the innate and adaptive immune systems protecting against various biological and chemical intruders. Overall, differentiated blood cells have a limited lifespan ranging from several hours (e.g. some granulocytes) up to several decades (e.g. memory T cells)⁶.

A malfunction of the blood system may arise by a cell type not forming e.g. by differentiation block or by a cell type not fulfil its function. This can lead to life-threatening disease or even death. In most of the cases this malfunction is caused by a genetic defect.

1.2 Primary immunodeficiency diseases

There are over 10.000 known monogenic diseases that are caused by single gene mutations. Although most of the monogenic disorders are relatively rare, they altogether affect about 0.5 - 1% of newborns and up to 10% of hospitalized patients^{7,8}. The best characterized monogenic diseases involve the hematopoietic system and include the primary immunodeficiency diseases (PIDs).

PIDs are caused by inherited mutations in genes required for the development and/or function of the immune system. Patients with PID lack an intact immune system resulting in increased susceptibility to infections, allergens, autoimmune reactions and cancer⁹. PIDs consist of over 200 different diseases which are all rare, chronic and usually fatal diseases¹⁰. Any component of the immune system can be affected a PID.

In some disorders, only a single part of the system is affected such as chronic granulomatous disease (CGD; also see chapter 1.7). In other diseases, there are multiple components can be affected such as in severe combined immunodeficiency (SCID)¹¹. Defects can be present in adaptive immune system (e.g. SCID and B-cell immunodeficiencies) or in the innate immune system (e.g. Toll-like Receptor (TLR), Natural Killer Cell (NK-cell) and myeloid differentiation primary response gene 88 (MYD88) deficiency)¹². Some of the best characterized PIDs are described in some more detail below.

Table 1: Selected examples of PIDs and their prevalence^{13,14}

Disease	Defect Gene	Mode of inheritance	Prevalence
Wiskott-Aldrich syndrome (WAS)	WASP	X-linked recessive	1/250,000
Selective IgA deficiency (SIgAD1)	Unknown	Autosomal dominant	1/600
X-linked agammaglobulinemia (XLA)	BTK	X-linked recessive	1/200,000
X-linked severe combined immunodeficiency (X-SCID)	IL2RG	X-linked recessive	1/50,000 to 1/100,000
Adenosine deaminase deficiency SCID	ADA	Autosomal recessive	1/200,000 to 1/1,000,000
JAK3-deficient SCID	JAK3	Autosomal recessive	<1/500,000
Leukocyte adhesion deficiency type I (LAD1)	ITGB2	Autosomal recessive	1/100,000
X-linked chronic granulomatous disease (X-CGD)	CYBB	X-linked recessive	1/250,000

Wiskott-Aldrich syndrome (WAS) is a X-chromosomal inherited disease caused by mutations in the WAS gene encoding the Wiskott-Aldrich syndrome protein (WASp), a cytoskeletal protein expressed by leukocytes and platelets¹⁴. WAS patients exhibit characteristic Immunologic abnormalities of white blood cells. Besides immunodeficiency, patients typically develop other symptoms like skin rashes (eczema), thrombocytopenia (low platelet counts), autoimmune reactions and malignant lymphoma. As a result, most of the affected patients have a lifespan of less than 10 years^{15–17}.

Severe combined immunodeficiency (SCID) summarizes more than 10 different genetic disorders characterized by defects in lymphoid development¹⁸. Without treatment, SCID patients die within the first year of life due to secondary infections. The most common form of the disease is X-linked SCID (X-SCID) which is caused by mutations in the interleukin-2 receptor gamma chain (IL2RG) gene located on the X-chromosome. Since the IL2 gamma chain is shared between receptors for IL-2, IL-4, IL-7, IL-9, IL-15 and IL-21 receptors, its inactivation results in a complete block of T-cells and NK-cell cell differentiation. Therefore, X-SCID patients have no

T-cells and NK cells and also lack functional B-cells due to the absence of CD4⁺ T-helper cells^{18,13}.

The second most common severe immunodeficiency disease is ADA-SCID, which caused by mutations in the adenosine deaminase (ADA) gene. ADA is a key enzyme in purine catabolism. Its absence leads to the accumulation of purine metabolites in blood plasma, which are toxic to lymphocytes (B, T and NK cells) and block their proliferation. Consequently, patients suffer from serious recurrent and life-threatening infections¹³. Symptomatic treatment of SCID aims to minimize bacterial, viral and fungal infections by providing sterile environments combined with antibiotic, antiviral and antifungal treatment.

Finally, an example for an immunodeficiency of myeloid cells is the chronic granulomatous disease (CGD). Granulocyte and macrophages of CGD patients fail to kill phagocytosed microorganisms due to an inherited defect of superoxide production caused by mutations in genes encoding for the NADPH oxidase complex^{19,20}. Since CGD was selected as a disease model gene therapy in this theses, it will be described in more detail in the section 1.7.

1.3 Molecular background of PIDs

A substantial fraction of hereditary monogenic blood disorders are caused by patient specific mutations dispersed over the entire locus of the affected gene²¹. The inheritance pattern of most PIDs is either X-linked recessive, autosomal recessive or rarely, autosomal dominant¹³. The mutation type of PID mutations includes point mutations, small deletions or insertions, large deletions, duplications, inversions and other more complex mutations^{22,23}.

PID mutations can affect any part of the locus: coding regions (exons), the promoter, regulatory regions, termination signals, splice donors / acceptors and also introns of the genes²². The most frequent PID mutations are in coding exons all resulting in protein dysfunction. Commonly there are point mutations which by nucleotide replacement create either a premature stop codon (nonsense mutation) or a new codon for an unrelated amino acid (missense mutation). Nonsense mutations are

usually sensed by the splicing machinery and trigger the nonsense mediated decay pathway (NMD) resulting in mRNA degradation. As a result, affected cells do not express the mutated protein²⁴. However, in some instances the mutated mRNAs escape NMD and are translated into truncated proteins. In either case the protein function is typically lost^{23,25}.

Another common type of PID are frameshift mutations. Frameshift mutations can involve small deletions or insertions (usually one or two nucleotides)²² which similar to nonsense mutations cause premature stop codons triggering NMD mediated mRNA degradation (see above). Figure 2 shows that approximately 77% of the PIDs (those shown in the Table 1) are caused by missense, nonsense and frameshift mutations and only a minority by more complex mutations.

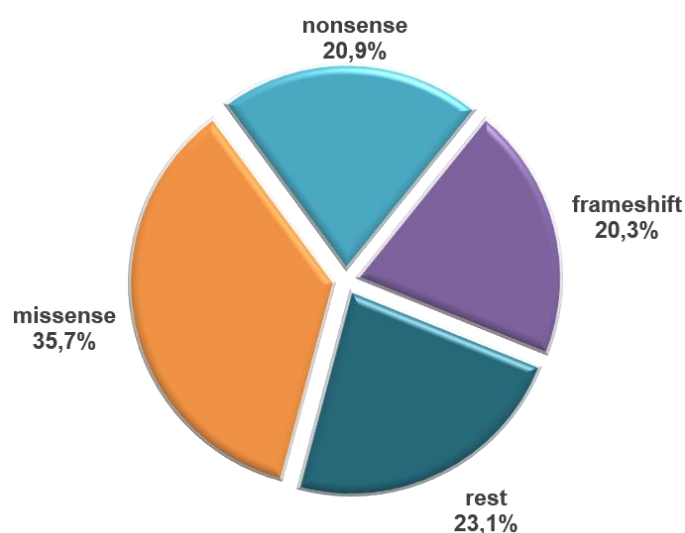


Figure 2: Mutation type of the selected examples of PIDs

Approximately 77% of all mutation are caused by point mutations (missense, nonsense and frameshift mutations). Only 23% of the mutations are more complex.

1.4 Hematopoietic stem cell transplantation and gene therapy

The standard PID treatment is allogeneic hematopoietic stem cell transplantation (HSCT). Based on the ability of HSCs to reconstitute the entire hematopoiesis, allogeneic HSCT has become the treatment of choice for a variety of congenital and acquired diseases, including PIDs and leukemias²⁶. However, successful allogeneic HSCT is largely dependent on the availability of Human Leukocyte Antigen matched (HLA-matched) donors. Transplantation of non- non-matched can cause life-

threatening graft-versus-host disease (GvHD) during which the engrafted lymphocytes elicit an immune response against host tissues perceived as foreign^{27,28}. However, 90% of PID patients with appropriate donors undergo complete immune reconstitution after HSCT and have a normal life expectancy²⁹. Unfortunately, for a significant number of patients HLA-matched donors cannot be found. This is particularly problematic in countries where comprehensive blood donor catalogs are unavailable³⁰.

The only alternative PID treatment strategy beside symptomatic treatment is gene therapy which involves transplantation of autologous HSCs after ex vivo gene transfer of a healthy copy of the compromised gene. As shown in Figure 3, gene therapy for PIDs includes (i) HSC isolation and purification, (ii) gene transfer by either transducing a functional gene copy or by in situ gene repair and (iii) reinfusion of the genetically manipulated HSC into the patient.

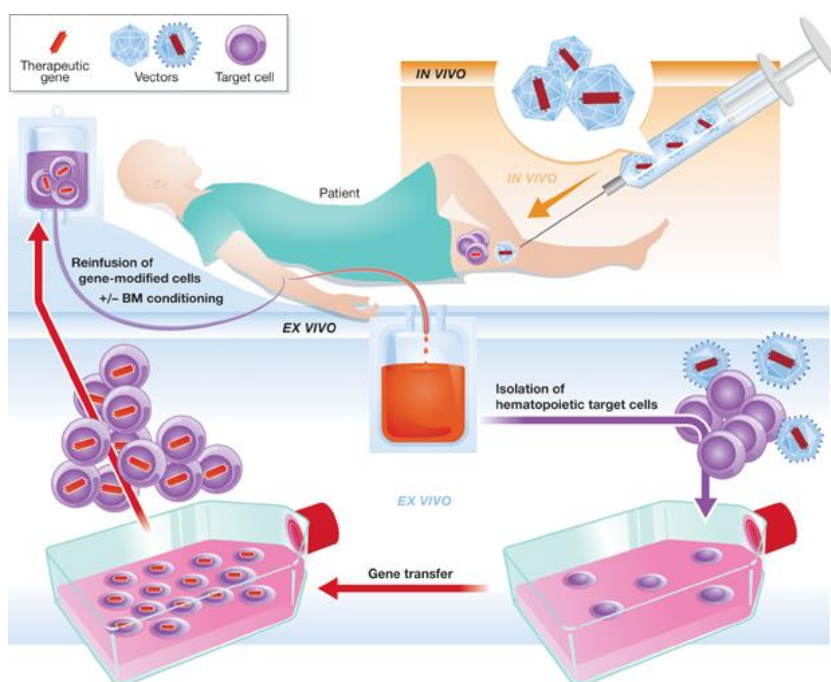


Figure 3: In vivo and ex vivo gene therapy concepts

For the in vivo treatment, the therapeutic gene is introduced directly into the body (e.g. muscle, liver) of the patient. For the ex vivo treatment, specific cells are first isolated from the patient body, modified in the laboratory with the appropriate vehicle and reinfused into the patient as an autologous transplant. BM, bone marrow. (Adopted from Kaufmann et al 2013⁹.)

The procedure circumvents GvHD and if successful, could be as lifesaving as allogeneic HSCT. The problems associated with gene therapy are discussed in more details in sections 1.5, 1.6 and 1.8.

1.5 Gene transfer vector systems

One of the greatest challenges of gene therapy is the efficient and safe transfer of gene correcting nucleic acids into the diseased cells. During the last 3 decades, several cell transduction systems have been developed, ranging from non-viral to viral-based vectors (Figure 4). Each system has its unique advantages and disadvantages, although for hematopoietic cells significant progress has been made with viral vectors to ensure reliable transgene expression with minimal adverse effects (see section 1.6)³¹. The disease itself determines the choice of the gene transfer vector.

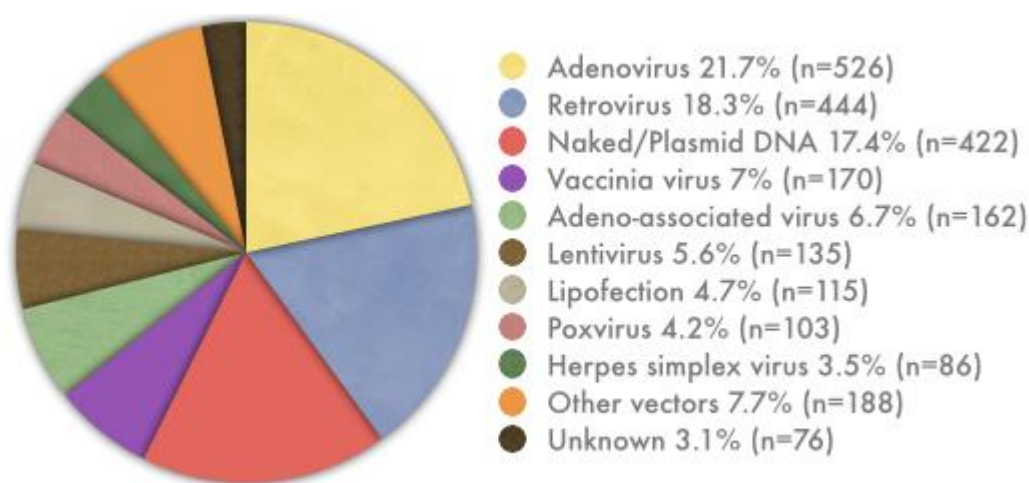


Figure 4: Overview of vectors used in gene therapy clinical trials

(Adopted from <http://www.abedia.com/wiley/images/1602vectors.jpg>)

1.5.1 Non-viral vectors

The prevalent non-viral vectors are expression plasmids consisting of transcription units expressing one or more genes of interest after introduction into the nucleus of a target cell. Plasmids are transduced into target cells by physical and/or chemical means. This process is collectively referred to as transfection. Figure 5 summarizes

the most common transfection methods. The most promising non-viral vectors delivered by transfection are liposomes and cationic polymers, which form a complex with DNA to form lipoplexes/polyplexes³². These nanomeric complexes are stable and enter cells usually by endocytosis³¹.

The major advantages of non-viral vectors are its simplicity of preparation, relatively low immunogenicity and by chemical low toxicity. Drawbacks are their relatively lower efficiency and that they have only terminate expression compared to the viral gene delivery^{31,33,34}.

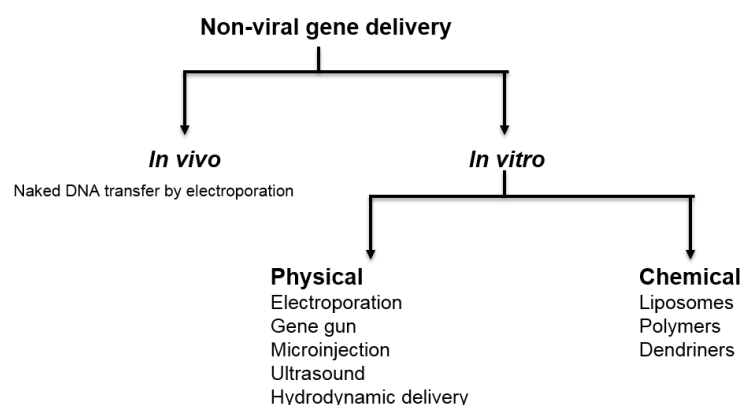


Figure 5: Various nonviral gene delivery systems (modified from Manjila et al. 2013)

An alternative non-viral gene delivery systems for long-term gene expression are transposons. Transposon systems are two component systems consisting of a transposon which is a genetic element flanked by inverted repeats (this genetic element could be any gene of interest) and an enzyme, the transposase, which catalyzes cutting and pasting of the transposon in and out of the genome^{35,36}. Unlike expression plasmids, transposons integrate stably into the genome and ensure long term expression of the inserted gene. Transposons are transduced into target cells together with a transposase-expressing plasmid usually by transfection (Figure 6)³⁷. Once inside the cell, the transposase triggers transposon integration which, depending on the transposon system, occurs more or less randomly throughout the genome (Figure 6). Transposons are naturally occurring mobile genetic elements, which are common to all organisms. The most popular transposons currently used are Sleeping Beauty (SB) and PiggyBac (PB)³⁵⁻³⁷. However, like the retroviral

vectors (see below) transposons are insertionally mutagenic and therefore associated with adverse effects^{38–42}.

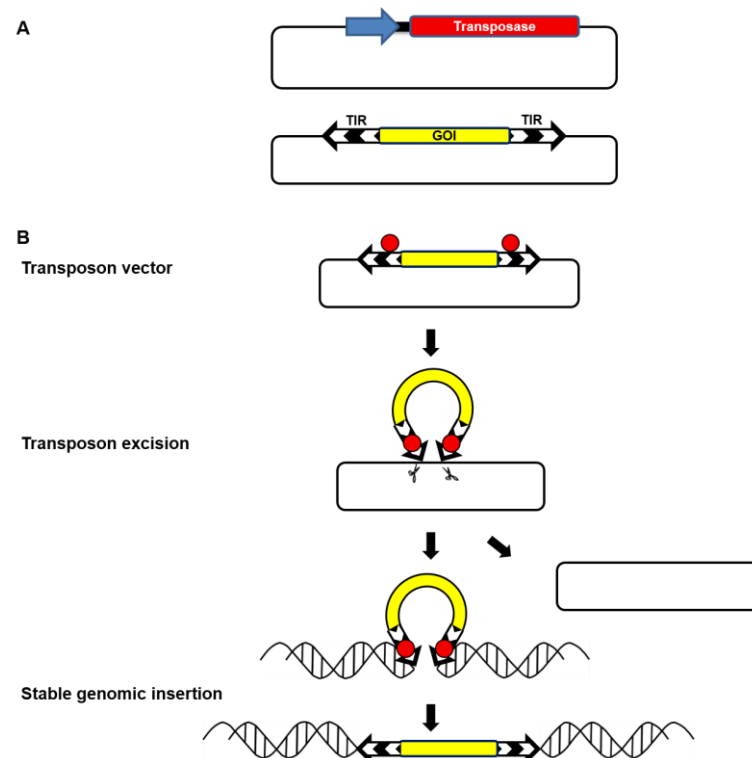


Figure 6: Transposon vector system for stable gene delivery (modified from Ivics et al. 2011³⁷)

Plasmid-based transposon vector system for transgene delivery. **A.** The vector system consists of two components: a plasmid containing a gene of interest (GOI) cloned between the transposon terminal inverted repeats (TIR; black arrows) and another plasmid expressing the transposase from a suitable promoter (black arrow). **B.** The transposase protein binds to the TIR of the transposon which carries the desired transgene, cleaves it from the donor plasmid and inserts the transposon into the genomic DNA.

1.5.2 Viral vectors

Viruses are natural nucleic acid transducers that enter susceptible cells by interacting with specific cell surface receptors. Once inside the cell, they release their genetic material which directs replication via engaging the cellular nucleic acid and protein synthesis machineries. In viral vectors, most of the viral genome is replaced by one or more genes of interest^{31,43}. Typically, the modified genomes are assembled into infectious particles in cell lines expressing the required viral proteins from separate expression plasmids⁴⁴. These cell lines are commonly referred to as viral producer or packaging cell lines.

Table 2: Viral vectors and their main properties

Vectors	γ-retrovirus	Lentivirus	Adenovirus	Adeno-associated virus
Genome	ssRNA	ssRNA	dsDNA	ssDNA
Packaging capacity	8.8 kb	9.6 kb	8.3 kb	4.8 kb
Target cells	Dividing cells	Dividing and Nondividing cells	Dividing and Nondividing cells	Dividing and Nondividing cells
Integration and transgene expression	Yes, Stable	Yes, Stable	No, Transient	No, Transient (Limited site-specific integration)
Drawbacks	Insertional mutagenesis, Infects only dividing cells	Insertional mutagenesis	Inflammation cytotoxicity, Short term expression	Low packaging capacity, Requires helper virus for replication

1.5.3 Non-integrating viral vectors

The most commonly used non-integrating viral vector are based on adenovirus. They have been used in 22% of all gene therapy trials including gene therapy of cancer and cystic fibrosis⁴⁵. Adenoviruses are double stranded DNA viruses that do not integrate their genome into the host genome. Instead, the adenoviral genome remains in the nucleus as an episomal element. Adenoviral vectors are relatively easy to produce at high concentrations and efficiently infect a large variety of both dividing and non-dividing cells^{44,46}. They express their cargo episomally and therefore, only transiently for 7-42 days. As 90% of all people have pre-existing antibodies to one or more adenovirus serotypes, the transgene expression can be further limited by an immune response eliminating the transduced cells⁴⁴. Therefore, it is advisable to determine anti-Adenoviruse antibody titers prior and during adenovirus-based gene therapy.

One alternative to adenoviruses, the adeno-associated viruses (AAV), demonstrate low immunogenicity. AAV are single-stranded DNA viruses that require a helper virus such as adenovirus or herpes simplex virus. AAV encodes for two proteins, rep and cap, which can be provided in-trans for virus production. Thus, to generate a recombinant AAV particle, rep and cap can be deleted from the viral genome to

leave only the virus inverted terminal repeats (ITRs)^{45,47}. The subtype AAV2 is the most commonly used vector. Although AAV belong to the group of non-integrating viruses, they still have limited genome insertion sites. An interesting feature of AAV is that they tend to stably integrate into the AAVS1 locus on chromosome 19⁴⁸. AAV vectors are currently considered the delivery tool of choice for in vivo therapy of inherited diseases in post-mitotic tissues. The major disadvantage of AAV vectors for gene delivery is their limited packaging capacity of up to 4.8 kb³¹. Another limitation is the frequently seen genomic integration of AAV genome into the host genome at other sites than AAVS1 on chromosome 19. Therefore, AAVs and AAV vectors are associated with tumorigenesis through insertional mutagenesis mainly into the proto-oncogenes^{48–50}

1.5.4 Integrating viral vectors

The most commonly used integrating viruses are γ -retroviruses and lentiviruses. γ -retroviruses (e.g. MLV) and lentiviruses (e.g. HIV) belong to the family of retroviridae, which are a positive-sense RNA viruses that replicate via a double stranded DNA intermediate⁴⁵. In infected cells the RNA genome is reversely transcribed into double-stranded DNA (dsDNA) which integrates randomly throughout the genome as a provirus. Reverse transcription and integration are catalyzed by the virally mRNA encoded enzymes reverse transcriptase and integrase. Proviruses are flanked by repetitive elements known as long terminal repeats (LTRs) that ensure viral replication by controlling mRNA transcription and processing^{45,47,51}

In retroviral vectors, all genes required for virus replication such as *gag* (encoding viral matrix, capsid, and nucleocapsid proteins), *pol* (encoding a *protease*, *reverse transcriptase*, and *integrase*), and *env* (encoding a bipartite membrane-anchored surface protein) are replaced by a gene of interest^{45,46,52}. To produce replication-defective infectious particles, the viral vector carrying the genes of interest is co-expressed with *gag*, *pol* and *env* encoding cassettes in dedicated packaging cell lines^{45,46,52} similar to those described above for adenovirus and AAV production.

1.5.5 Vectors of choice for gene replacement therapy in hematopoietic stem cells

Gene-replacement therapy in hematopoietic stem cells requires integrating vectors to ensure long term gene expression. Although more recently non-viral transposon vectors have been used^{39–41}, the vast majority of clinical trials employed γ -retroviral^{53,54} and lentiviral vectors^{55–57}. All integrating vectors insert semi-randomly throughout the genome, each having a specific integration pattern dictated by distinct features of different genomic regions^{39,58,59}. While γ -retroviral vectors (e.g. Moloney Murine Leukemia Virus (MMLV)) preferentially integrate in and around transcriptional start sites (TSS), including 5' regulatory elements such as promoters, enhancers, or locus control regions, lentiviral vectors more frequently integrate into actively transcribed genes^{39,59,60}. PiggyBac (PB) transposon replicates the integration pattern of MMLV whereas Sleeping Beauty (SB) transposon integration is arguably more random³⁹ (Figure 7).

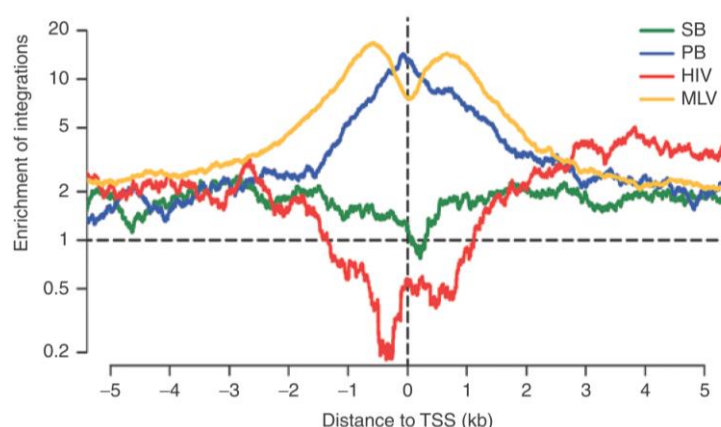


Figure 7: Analysis of integration frequencies around transcriptional start sites (TSS)

PB and MLV are enriched and HIV and SB are depleted around TSSs (adopted from Gogol-Döring et al. 2016³⁹).

However, in different γ -retroviral-based clinical trials, site integration induced deregulation of genes in the neighborhood and caused perturbation of the cellular phenotype. Insertions near oncogenes such as LMO2, MDS1-EVI1, PRDM16 or CCND2 led to clonal expansion eventually leading to leukemia. In each case leukemia was initiated by an LTR promoter/enhancer-triggered upregulation of

oncogene expression (Figure 8b)^{58,61}. In other instances, the therapeutic gene was silenced due to LTR promoter methylation^{15,58,62}.

Consequently, next generation gene therapy vectors lacking LTR promoter/enhancer elements were developed in which the therapeutic gene is expressed from an internal (maybe tissue-specific) promoter (Figure 8c). Clinical trials with these vectors commonly referred to as self-inactivating (SIN) vectors showed a substantially improved safety profile and less susceptibility to gene inactivation^{63,64}.

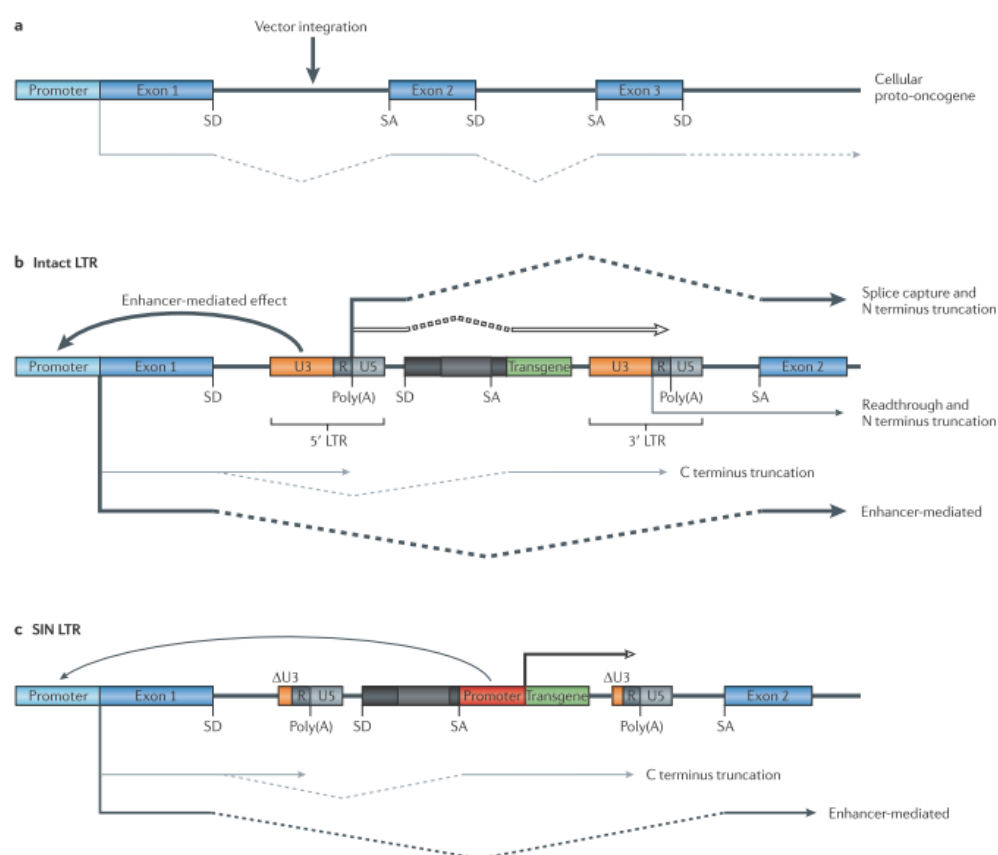


Figure 8: Mechanisms of the insertional mutagenesis mediated by the retroviral vectors

a) Vector integrations site is displayed in the first intron of a proto-oncogene. **b)** Y-retroviral vector integration in the first intron of a proto-oncogene could lead to upregulated transcription of the proto-oncogene from its cellular promoter by enhancer-mediated effects. Alternatively, expression driven by the 5' LTR or 3' LTR could give rise to a chimeric transcript an N-terminally truncated form of the oncogene. **c)** Self-inactivating (SIN) LTRs show improved are to be more safety and the expression of the therapeutic gene from an internal promoter is more robust (form Naldini L. 2011⁵⁸).

Currently, most clinical gene therapy trials employ third generation lentiviral vectors which are self-inactivating and usually rely on lineage-specific promoters^{9,65}. Although significantly less genotoxic than the earlier vector generations, the self-inactivating (SIN) lentiviral vectors are not entirely safe⁶⁶. They can still activate oncogenes from the internal promoter/enhancer elements as has been shown for Braf and Mak3k8 and inactivate tumor suppressor genes as has been shown for Pten and Rasa1⁶⁶. Thus, ideally gene replacement strategies would be substituted by designer endonuclease strategies capable of in situ gene correction (see below).

1.6 Clinical gene therapy trials for the treatment of PIDs

Over the last 3 decades, the curative potential of gene therapy has been demonstrated in many clinical phase I / II PID gene therapy trials^{67,68}. An overview of the most important past and ongoing PID gene therapy trials is given in Table 3.

There are many requirements for a successful gene therapy: a clear understanding of the molecular biology and genetics of the disease, the availability of tissue or cell targeted for the gene delivery, an effective therapeutic vehicle and an animal model that closely simulates the disease for preclinical studies.

Table 3: Clinical trials of HSC-based gene-therapy in PIDs

Disease	Therapeutic Gene	Total Patients	Vector/ Transcriptional Control Element	Method	Patient status and biological and clinical outcomes	Reference/ Clinical-trial identifier
WAS	WASP	3	LV / human endogenous WAS gene promoter	<i>ex vivo</i> gene transfer into CD34+	All patients AAW; stable engraftment with transduced cells; persistent clinical benefit and safety	¹⁶ NCT01515462
WAS	WASP	7	LV / human endogenous WAS gene promoter	<i>ex vivo</i> gene transfer into CD34+	6 patients AAW, 1 patient died of a pre-existing infection; stable engraftment with transduced cells; persistent clinical benefit and safety	⁶⁹ NCT01347242 NCT01347346 NCT02333760
X1-SCID	γ c	9	SIN- γ -RV / human EF1 α short	<i>ex vivo</i> gene transfer into CD34+	8 patients AAW, 1 patient died of an adenoviral infection; stable engraftment with transduced cells; persistent clinical benefit and safety in 7 patients	⁷⁰ NCT01410019 NCT01175239 NCT01129544
X1-SCID	γ c	10	γ -RV / MFG-LTR	<i>ex vivo</i> gene transfer into CD34+	4 of 10 developed T cell leukemia; Insertions near the LMO2 proto-oncogene; 1 patient died	^{71,72}
ADA-SCID	ADA	10	γ -RV / MLV-LTR	<i>ex vivo</i> gene transfer into CD34+	All patients AAW; stable engraftment with transduced cells; persistent clinical benefit and safety in 9 patients	⁷³ NCT00598481 NCT00599781
AR-CGD	NCF1	5	γ -RV / MLV-LTR	<i>ex vivo</i> gene transfer into CD34+	No clinical benefit	⁷⁴
X-CGD	CYBB	2	γ -RV / SFFV-LTR	<i>ex vivo</i> gene transfer into CD34+	Both patients developed functional neutrophils and developed MDS with monosomy 7 from insertional activation of MDS1-EVI1, PRDM16, SETBP1; 1 patient died	^{62,75} NCT00564759

AAW: alive and well

Taken together, these studies showed that clinical outcome is highly dependent on the transduction efficiency of hematopoietic cells capable of engrafting and functionally reconstituting the patient's hematopoiesis. However, clinical success was also dependent on the particular disease. While positive selection imposed on genetically reconstituted cells in patients with immunodeficiency diseases (e.g. X-SCID, ADA-SCID) increased the success rate, lack of such selection in patients with CGD was clearly detrimental because the non-modified, residual cells outcompeted the transplanted cells during hematopoietic regeneration^{9,76,77}.

1.7 Chronic granulomatous disease

Chronic granulomatous disease (CGD) is a rare inherited immunodeficiency affecting 1 in 250,000 individuals⁷⁸. CGD is characterized by the inability of phagocytes to eliminate ingested pathogens and is caused by mutations in any of the five genes of nicotinamide dinucleotide phosphate (NADPH) oxidase complex. In neutrophils, eosinophils, monocytes, and macrophages NADPH oxidase reduces molecular oxygen to reactive oxygen species (ROS) such as superoxide anion (O_2^-), hydrogen peroxide (H_2O_2), hypochlorite ion ($HOCl$) and hydroxyl radical (OH^\cdot), which kill the ingested microorganisms (Figure 9). Deficiencies of NADPH result in severe, recurrent and life-threatening bacterial and fungal infections such that affected patients require lifelong prophylactic treatment with antibiotics, antifungals and interferon gamma ($IFN-\gamma$)⁷⁸⁻⁸¹. To date, the only curative treatment for CGD is allogeneic bone marrow transplantation^{79,80,82} which is limited to patients having a suitable HLA-matched donor^{78,81,82}.

CGD is a genetically heterogeneous disease. Approximately, 60-70% of all CGD mutations affect the X-chromosomal *CYBB* gene encoding the protein gp91^{phox} (X-CGD). The other 30% have mutations in either the *CYBA*, *NCF1*, *NCF2* or *NCF4* genes encoding for p22^{phox}, p47^{phox}, p67^{phox}, and p40^{phox}, respectively. Unlike the X-linked *CYBB* mutations, these mutations are inherited in an autosomal recessive manner (AR-CGD)^{78,80,82,83}.

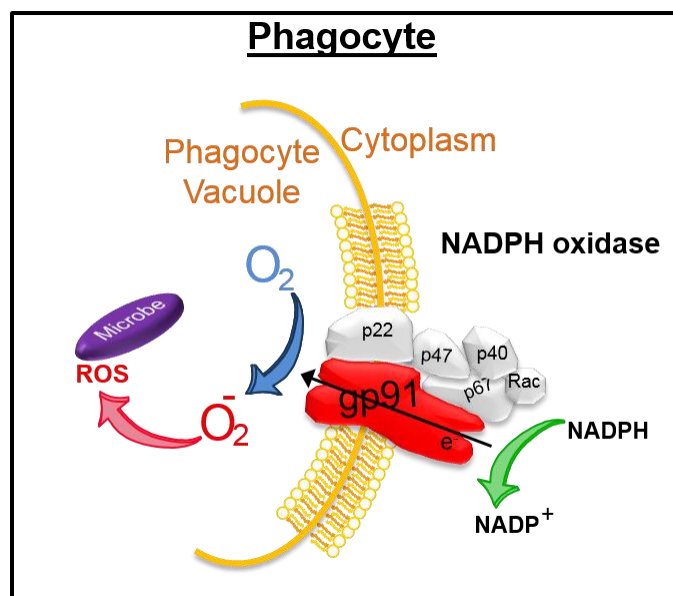


Figure 9: NADPH oxidase complexes in phagocytic cells

The NADPH oxidase comprises a membrane-bound gp91phox/p22phox heterodimer and other subunits (p67phox, p47phox, p40phox and Rac) which associates with this complex in the activated enzyme. The active NADPH produces reactive oxygen species (ROS) which are essential in defense against invading microbial pathogens.

1.8 Gene therapy for X-CGD

For X-CGD patients lacking a suitable donor, ex-vivo gene therapy combined with autologous bone marrow transplantation is the most promising therapeutic alternative. As has been shown previously in animal models and clinical trials, reconstitution of CYBB in only 5-10% of neutrophils is sufficient to cure the disease^{19,20,84}.

Several phase I/II gene therapy trials of X-CGD demonstrated the therapeutic potential transplantation of autologous HSCs retrovirally transduced ex vivo with retroviral vectors carrying healthy copies of CYBB for the treatment of X-CGD. The first X-CGD trial included five adult patients who received ex-vivo manipulated autologous HSC without myelosuppressive preconditioning resulting in less than 1% CYBB reconstitution⁸⁵. The second trial enrolled two adults (P1=26 years; P2=25 years) and was performed at Goethe-University Frankfurt. Unlike in the preceding trial, the patients received myelosuppressive preconditioning to enable the engraftment of genetically manipulated HSCs⁷⁵. Both trials employed γ -retroviral vectors expressing the *CYBB* cDNA from the LTR promoter. The Frankfurt patients

fared well for up to 5 months post-transplantation. In both patient CYBB function could be reconstituted in over 15% of granulocytes, leading to the eradication of preexisting life-threatening infections⁷⁵. However, after 5 months some CYBB-expressing cells underwent clonal expansion due to insertional activation of several oncogenes (i.e. PRDM16, MDS1/EVI1 and SETBP1). Although this temporarily increased the number of modified granulocytes from 15% to about 60%, CYBB expression was eventually silenced by LTR promoter methylation^{62,75,86}. Intriguingly, the LTR enhancer was not affected and continued to transactivate the nearby proto-oncogenes. In the end, EVI1 overexpression in both patients led to the genomic instability and preleukemic myelodysplasia with monosomy 7^{62,86}.

Similar genotoxicity was observed in an 8.5 year old boy treated in Zürich according to the Frankfurt protocol (Table 4). In all 3 further studies involving gamma-retroviral vectors and patient preconditioning the genetically manipulated cells failed to engraft (Table 4)⁸⁵.

Table 4: Summary of gene therapy trials for X-CGD including myelosuppressive strategies

(Modified form Grez et al. 2011⁸⁵.)

Center	Total Patients	Total conditioning	Vector/ Transcriptional Control Element	Significant engraftment >3 months	Genotoxicity	Reference
Frankfurt	2	Liposomal busulfan, 8.0 mg/kg	SF71gp91 ^{phox} (SFFV-LTR)	15% gene marking in CD15 ⁺ cells	Both patients developed clonal myeloproliferation and MDS with monosomy 7	62,75
Zürich	1	Liposomal busulfan, 8.8 mg/kg	SF71gp91 ^{phox} (SFFV-LTR)	20% gene marking in CD15 ⁺ cells	Development of clonal myeloproliferation	87
London	1	Melphalan, 140 mg/m ²	MFGS-gp91 ^{phox} (MLV-LTR)	No	No	
	3		SF71gp91 ^{phox} (SFFV-LTR)			
NIH	3	Busulfex, 10 mg/kg	MFGS-gp91 ^{phox} (MLV-LTR)	No	No	84
Seoul	2	Busulfex, 6.4 mg/kg + fludarabine, 120 mg/m ²	MFGS-gp91 ^{phox} (MLV-LTR)	No	No	88

Overall, the clinical X-CGD gene therapy revealed that CYBB function can be temporarily restored by *ex vivo* gene replacement although the toxicity of the employed strategy is prohibitively high, suggesting that significantly safer protocols are required, which guarantee a long-term and safe gene therapy.

1.9 Site-specific endonucleases

Over the last few decades, the tools available for genome manipulation have advanced significantly. Genome editing/engineering is widely used in the basic research to specifically *knock-in* or *knock-out* genes in given target cells (e.g. viruses, single cell organisms, plants, even vertebrates). Furthermore it is used for development of novel therapies. A wide array of powerful gene editing systems are now available, which are called engineered or designer endonucleases^{89,90}. Designer endonucleases are engineered enzymes that can introduce DNA double strand breaks (DSBs) with high specificity into the desired target sequences. To be useful for genome engineering, the endonucleases must have two particular functions. First, there must be a specific recognition domain (protein or RNA) which recognizes long target sequences (ideally, long enough for unique occurrence in the eukaryotic genome). Second, they must cleave in the targeted sequence. To date, four types of endonucleases exist; (1) meganucleases, (2) zinc finger nucleases (ZFNs), (3) transcription activator–like effector nucleases (TALENs) and (4) CRISPR-associated nuclease Cas9.

1.9.1 Meganucleases

Meganucleases (also called homing endonucleases) are natural endodeoxyribonucleases characterized by a large target recognition sites (14-40 bp)^{91,92} which generally occur only once in any given genome. They are found in bacteria, archaea, phages, protists, fungi, and plants^{93,94}. Although the isolation of natural meganucleases with new target specificities is tedious and slow, target specificity can be modified in existing meganucleases by protein engineering^{92,93}. However, such modifications are challenging^{92,93}.

1.9.2 Zinc-finger nucleases

Zinc-finger nucleases (ZFNs) are chimeric endonucleases, consisting of individual zinc-finger protein (ZFP) DNA binding motifs fused to the cleavage domain of the *FokI* restriction endonuclease⁹⁵. ZFNs typically contain between 3 – 6 ZFP motifs. Each motif recognizes 3 – 4 bp, which results in a recognition site of typically 9 – 18 bp^{90,96}. As the *FokI* cleavage domain acts only in form of a dimer in order to cleave DNA and therefore a pair of ZFNs targeting the sense and antisense strand are required^{90,97}. The main drawback of the ZFNs is the limited number of available ZFP motifs and that the individual ZFP domains do not independently bind to their specific DNA element but rather influence the binding of the neighboring domain, which results in many ZFNs leading to off-target effects. This cooperative binding makes the ZFN design difficult and time consuming as it requires sophisticated protein engineering and specialized methodology, thus preventing their widespread use^{90,98}.

1.9.3 Transcription activator-like effectors nucleases

Transcription activator-like effectors nucleases (TALENs) are, similar to ZFNs, chimeric endonucleases in which a transcription activator-like effector (TALE) DNA binding domain is fused to the DNA cleavage domain of *FokI*^{99,100}. The DNA binding domains consist of highly conserved 33-35 amino acid sequence in which only the amino acid located at the position 12 and 13 vary. These two positions, referred to as the Repeat Variable Di-residue (RVD), are highly variable and responsible for specific base pair recognition^{90,99}. Naturally, TAL effector proteins are secreted by *Xanthomonas* bacteria during plant infection resulting in the modulation of gene expression¹⁰⁰. In contrast to ZFN the individual domains do not influence binding of the neighboring domain, making TALENs design and assembly is relatively simple as it does not require extensive screening for target specificity^{99–102}. Central problems with TALENs are, that they are sensitive to cytosine methylation, especially at CpG region and many TALEN pairs provide little or no mutagenesis activity.

1.9.4 Clustered regularly interspaced short palindromic repeats/Cas9

Clustered regularly interspaced short palindromic repeats (CRISPR) are defining components of the genomes of most bacteria and archaea and are part of their adaptive immune system defending them against phage and plasmid DNA infection¹⁰³. The first report on the CRISPR array was from 1987 by Ishino and colleagues, who found in *E.coli* 29 nt repeats that were interspersed by 32 nt long non-repetitive sequences (spacers)¹⁰⁴. Three different types of CRISPR/Cas9 systems have been characterized thus far. The most widely used CRISPR/Cas9 system of *Streptococcus pyogenes* (spCas9/sgRNA) consists of three components: the CRISPR-associated DNA cleaving endonuclease Cas9 protein (~160 kDa, ~4.2 kb), a target DNA sequence recognizing RNA which is transcribed from short DNA sequences known as protospacers that are separated by short palindromic sequences clustered in the bacterial genome in the CRISPR array, (crRNA), and a trans-activating crRNA (tracrRNA) required for crRNA transcription^{90,105}. For genome editing, the crRNA and tracrRNA were fused into a fully functional single guide RNA (sgRNA) of 110 nt¹⁰⁵. Additionally, functional spCas9 requires a so called protospacer adjacent motif (PAM) sequence (i.e. 5'NGG) located downstream of the sgRNA target sequence which determines the exact position of DNA cleavage occurring always 3 nt upstream of the PAM¹⁰⁶. Cas9 consists of the HNH- and the RuvC-like nuclease domains, which cleave the coding and the non-coding DNA strands, respectively. Unlike the other designer endonucleases, CRISPR/Cas9 is addressed to the target site solely by a RNA molecule, without any protein engineering^{107–109}. Since the CRISPR/Cas9 system relies on RNA/DNA base-pairing it circumvents problems encountered with the other, protein based editing systems such as inactivating methylations and does not require protein engineering. CRISPR/Cas9 RNA-guided nucleases (RGNs) are easy to make and can be applied to a large variety of genome editing tasks including targeted gene activation, multiple gene targeting (multiplexing) and epigenetic manipulation^{110–112}.

1.10 Molecular outcomes of genome editing

DSB generated by site-specific designer endonuclease activate the cellular repair machinery which restores the lesion either by homology directed repair (HDR) or by non-homologous end joining (NHEJ). HDR is an error-free process because it requires a donor DNA sequence as a repair template^{90,111}. Thus, HDR can be exploited to create specific sequence changes, including the targeted addition of whole genes (knock-ins)^{90,113,114}. In contrast, NHEJ restores DSBs in absence of a template by religating the DNA ends – a process associated with random nucleotide insertions or deletions (indels)^{109,114}. Because indels frequently cause mutations, designer endonuclease systems and in particular the CRISPR/Cas9 system are being employed extensively for functional genetic screens^{115–117} and for the accelerated production of knock-out animals^{118–121}. However, designer endonucleases also have the potential to correct genetic mutations directly in affected tissues and cells to treat diseases that are not curable by traditional therapies (Table 5)^{97,111,122–124}.

Table 5: Examples of applications of genome editing to therapeutic model disease.

(Adopted from Cox et al. 2015¹¹¹)

Disease type	Nuclease platform	Therapeutic strategy	Reference
Hemophilia B	ZFN	HDR-mediated insertion of correct gene sequence	125
HIV	ZFN and CRISPR	NHEJ-mediated inactivation of CCR5	126–129
Duchenne muscular dystrophy (DMD)	TALEN and CRISPR	NHEJ-mediated removal of stop codon HDR-mediated gene correction	130,131
Hepatitis B virus (HBV)	TALEN and CRISPR	NHEJ-mediated depletion of viral DNA HDR-mediated	132,133
SCID	ZFN	HDR-mediated insertion of correct gene sequence	134
Cataracts	CRISPR	HDR-mediated correction of mutation in mouse zygote	124
Cystic fibrosis Hereditary	CRISPR	HDR-mediated correction of CFTR in intestinal stem cells	135
Hereditary tyrosinemia	CRISPR	HDR-mediated correction of mutation in liver	136

1.11 Aim of this work

A large part of hereditary monogenic disorders affecting the hematopoietic system, are caused by patient-specific mutations spread over the entire locus of the affected gene. These patients often lack an intact immune system resulting in increased susceptibility to infections, allergens, autoimmune reactions and cancer. Designer endonucleases, especially CRISPR/Cas9, hold the potential to significantly improve the results of future clinical personalized gene therapy approaches for these patients. Hitherto, the correction of monogenic disorders using designer endonucleases are primarily based on homology directed repair (HDR), which is dependent on an exogenous DNA template. The efficiency of the gene correction by means of HDR is generally very low, reduced even further by the cell cycle dependence of HDR. Additionally, in hematopoietic stem- and progenitor cells (HSPC) the dominant DSB-repair pathway is the NHEJ pathway.

The aim of this thesis was to develop a template-free NHEJ-mediated DNA repair strategy for personalized gene therapy of primary immunodeficiency diseases (PIDs) using site-specific designer endonucleases. In theory, approximately one third of the indels associated with NHEJ should restore the open reading frame (ORF) disrupted by a particular disorder point mutation. Depending on the location and type of the point mutation, some of the reconstituted ORFs should either completely or partially recover protein function. The following objectives were tackled:

- Comparing the efficiency of the DSB repair mechanism of HDR vs. NHEJ using site-specific designer endonucleases
- Restore a point mutated Hprt gene by NHEJ to reconstruction the non-mutated gene
- Generate a restore a point mutated EGFP reporter
- Generate cell lines expressing a single copy CYBB reporter genes harboring patient specific point mutations
- Test the efficiency of CRISPR/Cas9-mediated gene correction delivered by integrase-defective lentiviral (IDLV) into the mutant CYBB target cells
- Estimate the target mutation efficiency at the endogenous CYBB locus in human hematopoietic cells

2. Materials and Methods

2.1 Material

2.1.1 Software and online tools

Name	Manufacturer
ApE-A plasmid Editor	by Mikel Wayne Davis (University of Utah)
Benchling (CRISPR predictor)	https://benchling.com/
CCTop - CRISPR/Cas9 (CRISPR predictor)	http://crispr.cos.uni-heidelberg.de/
Clustal Omega (Multiple Sequence Alignment)	http://www.ebi.ac.uk/Tools/msa/clustalo/
CYBBbase variation browser	http://structure.bmc.lu.se/idbase/CYBBbase/browser.php?content=browser
Diva7 FACS-software	Becton Dickinson, Franklin Lakes, USA
Ensembl	http://www.ensembl.org/index.html
Flowing software 2.5.1	http://flowingsoftware.btk.fi/
GraphPad Prism 5	GraphPad software, La Jolla, USA
Lasergene 7	DNASTAR, Madison, USA
Mendeley Desktop	Elsevier, Amsterdam, Netherlands
Microsoft Office 2013	Microsoft, Redmond, USA
Oligo Tm Analyzer	http://tmcalculator.neb.com/#/
Quantity One 1-D analysis software	Bio-Rad, Hercules, USA
SMART modular architecture research tool	http://smart.embl-heidelberg.de/

2.1.2 Hardware and consumables

Device	Manufacturer
Analytical balance	Sartorius, Göttingen, Deutschland
Agarose gel electrophoresis power supply - EPS 500/400	Pharmacia LKB, Uppsala, Sweden
Agarose gel running chamber - 40-0911	Peqlab Biotechnologie, Erlangen, Germany
Bacteria shaker Roth KS 15A	Roth, Karlsruhe, Germany
Centrifuge - Rotina 35	Hettich, Tuttlingen, Germany
Centrifuge – Rotanta 460	Hettich, Tuttlingen, Germany
Centrifuge - Microcentrifuges 5424	Eppendorf, Hamburg, Germany
BD FACSAria TM III	Becton Dickinson, Franklin Lakes, USA
BD LSRFortessa TM	Becton Dickinson, Franklin Lakes, USA
Gel documentation system - E.A.S.Y.RH-3	Herolab, Wiesloch, Germany
incubator Heraeus B6030 (Bacteria)	Thermo Fisher Scientific, Massachusetts, USA

Incubator Heracell 150i	Thermo Fisher Scientific, Massachusetts, USA
Microscope Zeiss ID03	Zeiss, Jena, Germany
NanoDrop 2000c	Thermo Fisher Scientific, Massachusetts, USA
Nucleofector™ 2b	Lonza, Basel, Switzerland
Odyssey FC	LI-COR Bioscience, Bad Homburg, Germany
PAGE Power supply - Power Pac 1000	Bio-Rad, Hercules, USA
PAGE system - XCELL SureLock™	Invitrogen, Carlsbad, USA
Pipette	Gilson, Middleton, USA
PCR device - Mastercycler® pro	Eppendorf, Hamburg, Germany
PCR device – T100™ Thermal Cycler	Bio-Rad, Hercules, USA
pH-meter - Hanna pH 210	Hanna instruments, Woonsocket, USA
Photometer	Eppendorf, Hamburg, Germany
Quadro-MACS cell separator	Miltenyi Biotech, Bergisch Gladbach, Germany
Sonifier – VW 2070	Bandelin electronic, Berlin, Germany
Laminar flow; Sterile bank Class II Nuaire	Zapf Instruments, Sarsted, Germany
Table centrifuge - Himac CT15RE	Hitachi, Tokyo, Japan
Thermocycler - Mastercycler ^R Pro	Eppendorf, Hamburg, Germany
Wet Transfer blot module - XCELL II™ Blot Module	Invitrogen, Carlsbad, USA
Ultracentrifuge - Optima™ XPN 80K	Beckman Coulter, Brea, USA
Vortexer VF-2 Janke&Kunkel	IKA, Staufen, Germany
Water bath GFL1083	GFL, Burgwedel, Germany

2.1.3 Plastic ware and consumables

Item	Manufacturer
Cell culture flasks, plates and dishes	Greiner bio-one, Kremsmünster, Austria
Filter Tips	StarLab, Hamburg, Germany
Flow cytometry tubes	Sarstedt, Nümbrecht, Germany
Tips	StarLab, Hamburg, Germany
Mikrofilter (0,22µm, 045µm)	Merck Millipore, Darmstadt, Germany
1.5 – 2 ml reaction tubes	Sarstedt, Nümbrecht, Germany
15 – 50 ml reaction tubes	Sarstedt, Nümbrecht, Germany
PCR tubes	Greiner bio-one, Kremsmünster, Austria
Ultracentrifuge tubes	Beckman Coulter, Brea, USA

2.1.4 Chemicals and reagents

Chemical	Manufacturer
1 kB plus DNA ladder	NEB, Frankfurt, Germany
2-Propanol	Sigma-Aldrich, Taufkirchen, Germany
10x Restriction buffer 1,2,3,4	New England Biolabs, Frankfurt, Germany
Acrylamide 40%	Amresco, Solon, USA
Agarose	Roth, Karlsruhe, Germany
Ampicillin sodium salt	Roth, Karlsruhe, Germany
APS 10%	Roth, Karlsruhe, Germany
Antarctic Phosphatase Buffer 10x	NEB, Frankfurt, Germany
Bovine Serum Albumine (BSA)	Applichem, Darmstadt, Germany
Chloroform	Fluka, Deisenhofen, Germany
Chloroquine	Sigma-Aldrich, Taufkirchen, Germany
DMEM	Gibco, Karlsruhe, Germany
dNTP mix (10 µM each)	NEB, Frankfurt, Germany
Dihydrorhodamine123	Sigma-Aldrich, Taufkirchen, Germany
Dimethylsulfoxide (DMSO)	Sigma-Aldrich, Taufkirchen, Germany
Dithiothreitol (DTT)	AppliChem, Darmstadt, Germany
ddH ₂ O	Sigma-Aldrich, Taufkirchen, Germany
Ethylenediaminetetraacetic acid (EDTA)	Roth, Karlsruhe, Germany
Ethanol	Sigma-Aldrich, Taufkirchen, Germany
Ethidiumbromide	Roth, Karlsruhe, Germany
Fetal calf serum (FCS)	Sigma-Aldrich, Taufkirchen, Germany
Formaldehyde	Merck, Darmstadt, Germany
Glycerol	Roth, Karlsruhe, Germany
Glycine	Merck, Darmstadt, Germany
Hydrochloric acid (HCl)	Roth, Karlsruhe, Germany
HDGreen Plus	Intas, Göttingen, Germany
HEPES	Roth, Karlsruhe, Germany
Isopropanol	Fluka, Deisenhofen, Germany
LB-Agar	Roth, Karlsruhe, Germany
LB-Medium	Roth, Karlsruhe, Germany
L-glutamine	PAA Laboratories, Pasching, Austria
Methanol	Merck, Darmstadt, Germany
Methylene blue	Sigma-Aldrich, Taufkirchen, Germany
Sodium chloride (NaCl)	Roth, Karlsruhe, Germany
Sodium hydroxide (NaOH)	Roth, Karlsruhe, Germany

NBT	Sigma-Aldrich, Taufkirchen, Germany
Paraformaldehyde (PFA)	Sigma-Aldrich, Taufkirchen, Germany
PBS (w/o Ca ²⁺ Mg ²⁺)	PAA Laboratories, Pasching, Austria
Penicillin/Streptomycin (100x)	PAA Laboratories, Pasching, Austria
PMA	Sigma-Aldrich, Taufkirchen, Germany
Polyethylenimin (PEI)	Sigma-Aldrich, Taufkirchen, Germany
Polybrene	Sigma-Aldrich, Taufkirchen, Germany
Ponceau S	AppliChem, Darmstadt, Germany
Precision Protein Standard	Biorad, Richmond, USA
Q5 ® Polymerase Buffer 5X	NEB, Frankfurt, Germany
Red Taq Buffer 10x	Sigma-Aldrich, Taufkirchen, Germany
Restriction Buffers 10x	NEB, Frankfurt, Germany
RPMI	Gibco, Karlsruhe, Germany
Sucrose	Roth, Karlsruhe, Germany
SuperSignal® West Femto Luminol/ Enhancer Solution	Thermo Scientific, Rockford, USA
SuperSignal® West Femto Stable Peroxide Solution	Thermo Scientific, Rockford, USA
SuperSignal® West Pico Luminol/ Enhancer Solution	Thermo Scientific, Rockford, USA
SuperSignal® West Pico Stable Peroxide Solution	Thermo Scientific, Rockford, USA
T4 DNA ligase Buffer	NEB, Frankfurt, Germany
TEMED	Roth, Karlsruhe, Germany
Triton X-100	Sigma-Aldrich, Taufkirchen, Germany
Trypan blue	Sigma-Aldrich, Taufkirchen, Germany
Tween-20	Roth, Karlsruhe, Germany
Tris	Roth, Karlsruhe, Germany
X-Gal	Sigma-Aldrich, Taufkirchen, Germany
beta-Mercaptoethanol	Fluka, Deisenhofen, Germany

2.1.5 Buffers and solutions

Buffer	Composition
Blocking buffer (WB)	5% non-fat milk powder in 1X TBS-T
Bis-Tris buffer	1.25 M Bis-Tris pH 6.5 in H ₂ O
DNA-loading Buffer (10x)	100 mg Bromophenol blue 100 mg Xylen cyanol 33 ml 150 mM Tris pH 7.6 60 ml Glycerol 7 ml H ₂ O
FACS fixation solution	2% formaldehyde in PBS
LB Agar	35 g agar powder mix in 1 l H ₂ O
LB medium	20 g LB broth mix in 1 l H ₂ O
MACS buffer	0.5% BSA 2 mM EDTA in PBS pH 7.2
SDS cell-lysis buffer for gDNA	50 mM Tris-HCl pH 8 100 mM NaCl 1 mM EDTA pH 8 0.5% SDS
SDS cell-lysis buffer for protein	100 mM Tris-HCl pH 8 150 mM NaCl 10 mM EDTA pH 8 10% SDS
SDS sample buffer (4X)	250 mM Tris-HCl pH 6.8 8% SDS 40% glycerol 0.2% bromphenolblue
Sorting buffer	2% FCS 1% Pen/Strep 1 mM EDTA in PBS
TAE buffer (50X)	242 g Tris 57.1 ml Acetic acid 100 ml EDTA (0.5M) pH 8.0
TBS-T	100 mM Tris-HCl pH 7.4 150 mM NaCl 0.05% Tween-20
Western Blot lysis buffer	50 mM Tris pH 7.4, 0.15 M NaCl, 2 mM EDTA, 1% NP-40 in H ₂ O
20% sucrose	100 g sucrose 400 ml PBS

2.1.6 Enzymes

Enzyme	Manufacturer
Antarctic Phosphatase	NEB, Frankfurt, Germany
RedTaq Polymerase	Sigma-Aldrich, Taufkirchen, Germany
Proteinase K	Sigma-Aldrich, Taufkirchen, Germany
Restriction Endonucleases	NEB, Frankfurt, Germany
Q5® High-Fidelity DNA Polymerase	NEB, Frankfurt, Germany
T4 DNA ligase	NEB, Frankfurt, Germany
Trypsin-EDTA (10x)	Sigma-Aldrich, Taufkirchen, Germany

2.1.7 Commercially available kits

Name	Application	Manufacturer
Amaxa® Cell Line Nucleofector Kit® V	DNA Nucleofection in cells	Lonza, Basel, Switzerland
DNA Clean & Concentrator™ - 5	DNA cleaning	Zymo Research, Irvine, USA
Human CD34 isolation kit	Cell isolation	Miltenyi Biotech, Bergisch Gladbach, Germany
Nucleobond PC 500	Plasmid preparation (Maxi)	Macherey-Nagel, Düren, Germany
TOPO-TA cloning kit	PCR-fragment cloning	Invitrogen, Carlsbad, USA
Zymoclean™ Gel DNA Recovery Kit	Gel extraction	Zymo Research, Irvine, USA
Surveyor® Mutation Detection Kit	Detect mutations in DNA	IDT, Coralville, USA
p24 ELISA kit (INNOTEST)	Viral titer determination	Fujirebio, Hannover, Germany

2.1.8 FACS antibodies

Antigen	Fluorophore	Target spec.	Manufacturer
CD11b	APC	Human	Becton Dickinson, Franklin Lakes, USA
CD34	PE / APC-Cy7	Human	Miltenyi Biotech, Bergisch Gladbach, Germany
CYBB(7D5)	FITC	Human	MBL, Nagano, Japan

2.1.9 Western blot antibodies

Antigen	Working dilution	Species	Target spec.	Manufacturer
EGFP	1:1000	mouse		Roche, Basel, Switzerland
GAPDH	1:1000	rabbit	human	Cell Signaling, Frankfurt am Main, Germany
CYBB (moAB48)	1:500	mouse	human	LifeSpan BioSciences, Seattle, WA, USA

Note: All the antibodies were diluted in 5% low-fat milk.

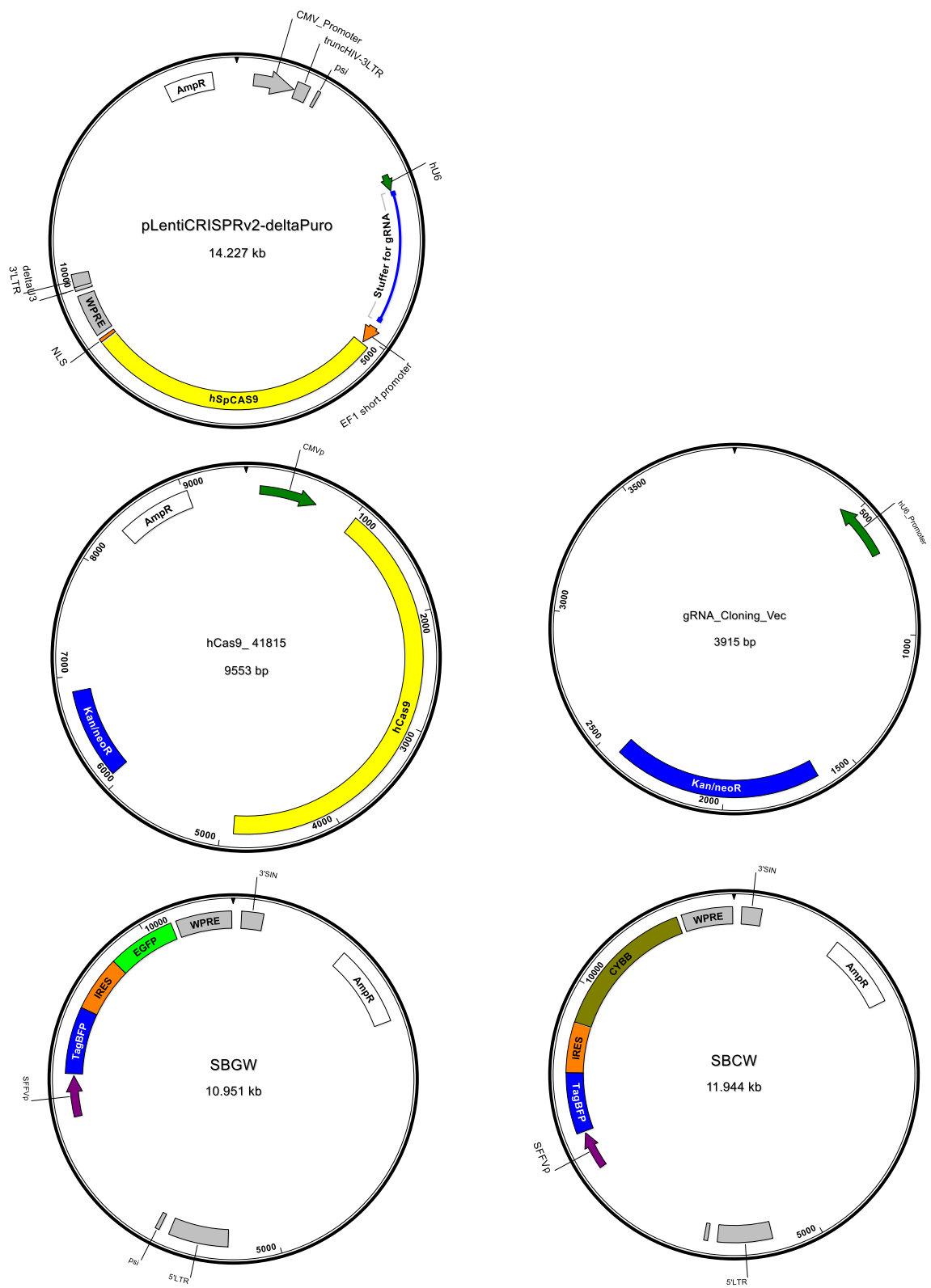
2.1.10 Cytokines

Cytokine	Manufacturer
Human Flt-3	Peprtech, New Jersey, USA
Human G-CSF	Peprtech, New Jersey, USA
Human IL-3	Peprtech, New Jersey, USA
Human SCF	Peprtech, New Jersey, USA
Human TPO	Peprtech, New Jersey, USA

2.1.11 Plasmids

Name	Function	Source
psPAX2	HIV-1 packaging plasmid encoding for gag, pol, rev und tat under control of the CMV promoter	Addgene (ID: 12260)
psPAX2-D64V	HIV-1 packaging plasmid encoding for gag, pol, rev und tat under control of the CMV promoter	Addgene (ID:63586)
pMD2.G	Packaging plasmid expressing the VSV-envelope under control of the CMV promoter	Addgene (ID:12259)
pCAG-T7-TALEN1 HPRT	Genome editing	Obtained from R. Kuhn
pCAG-T7-TALEN2 HPRT	Genome editing	Obtained from R. Kuhn
pmaxGFP	Nucleofection control vector encoding GFP under control of the CMV promoter	Lonza, Cologne, Germany
pHPRT-P2A-EGFP-pA	Donor vector for targeting HPRT locus	This thesis
pHR ^{SIN} -SEW	HIV-1 transfer-vector encoding EGFP under control of the SFFV promoter, contains WPRE, SIN-configuration	Obtained from M. Grez Demaison et al. 2002 ¹³⁷

pHR'SIN-SBW	HIV-1 transfer-vector encoding TagBFP (developed from TagRFP ¹³⁸) under control of the SFFV promoter, contains WPRE, SIN-configuration	Obtained from M. Grez
pHR'SIN-SBGW	Same configuration as pHR'SIN-SBW, encoding TagBFP-IRES-EGFP	This thesis
pHR'SIN-SBmGW	Same configuration as pHR'SIN-SBW, encoding TagBFP-IRES-mutEGFP	This thesis
pHR'SIN-SBCW	Same configuration as pHR'SIN-SBW, encoding TagBFP-IRES-CYBB	This thesis
pHR'SIN-SB45CW	Same configuration as pHR'SIN-SBW, encoding TagBFP-IRES-45CYBB	This thesis
pHR'SIN-SB54CW	Same configuration as pHR'SIN-SBW, encoding TagBFP-IRES-54CYBB	This thesis
pHR'SIN-SB134CW	Same configuration as pHR'SIN-SBW, encoding TagBFP-IRES-134CYBB	This thesis
pHR'SIN-SB173CW	Same configuration as pHR'SIN-SBW, encoding TagBFP-IRES-173CYBB	This thesis
phCas9	Expresses human codon optimized Cas9 nuclease for genome engineering under control of the CMV promoter	Addgene (ID:41815)
pgRNA Cloning Vector	An empty gRNA expression vector, used to create a gRNA to a specific sequence under control of the hU6 promoter	Addgene (ID:41824)
pLentiCRISPRv2	contains two expression cassettes, hSpCas9 + Puro under control of the EF1 α -short promoter and the chimeric guide RNA under control of the hU6 promoter	Addgene (ID:52961)
pLentiCRISPRv2 Δ Puro	Same configuration as pLentiCRISPRv2 delta-Puro	This thesis



2.1.12 Cell lines and culture media

Name	Genotype	Media	Source
HEK293T	Human embryonic kidney cell line genetically engineered to express the large T antigen	DMEM 10% FCS 1% L-Glutamine 1% Pen/Strep	DSMZ: ACC 635
32D	Mouse	RPMI 1640 10% FCS 1% L-Glutamine 1% Pen/Strep 10% (v/v) WEHI3B conditioned media	DSMZ: ACC 411
PLB985	Human acute myeloid leukemia cell line (derivative of HL-60)	RPMI 1640 10% FCS 1% L-Glutamine 1% Pen/Strep	M. Grez DSMZ: ACC 139
XCGD-PLB985	Human acute myeloid leukemia cell line with a disrupted CYBB gene by homologous recombination	RPMI 1640 10% FCS 1% L-Glutamine 1% Pen/Strep	M. Grez
WEHi-3B	Mouse myelomonocytic leukemia cell line to produce iL3 cytokine	Iscove basal Medium 10% FCS 1% L-Glutamine 1% Pen/Strep 0.1% beta-Mercaptoethanol	DSMZ: ACC 26

2.1.13 Bacterial stains

Name	Genotype	Source
DH5 α	<i>F- ϕ80lacZΔM15 Δ(lacZYA-argF)U169 recA1 endA1 hsdR17(rK-,mK+) phoA supE44 thi-1 gyrA96 relA1</i>	
TOP 10 <i>E. coli</i>	<i>F- mcrA Δ(mrr-hsdRMS-mcrBC) ϕ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara-leu)7697 galU galK rpsL (Str^R) endA1 nupG</i>	Invitrogen

2.1.14 Oligonucleotides

All oligonucleotides were ordered from Sigma Aldrich

Oligo		Sequence (5'-3')
Cloning of EGFP sgRNAs		
BH001	g32DG1 s	caccgGAGGAAGTTCACCGGCG
BH002	g32DG1 as	aaacGTGAACAGTTCCTCGCCCTc
BH003	g32DG2 s	caccgGAGGAAGTTCACCGGCG
BH004	g32DG2 as	aaacCGCCGGTGAACAGTTCCTCc
BH037	g32DG1.4 s	caccgGGGCGAGGAAGTTCAC
BH038	g32DG1.4 as	aaacGTGAAACAGTTCCTCGCCCCc
BH039	g32DG1.7 s	caccgAAGGGCGAGGAAGTTCAC
BH040	g32DG1.7 as	aaacGTGACAGTTCCTCGCCCTTc
BH041	g32DG2.3 s	caccgGGAAGTTCACCGCGGCG
BH042	g32DG2.3 as	aaacCGCCGCGGTGAACAGTTCc
BH043	g32DG2.5 s	caccgCGAGGAAGTTCACGGCG
BH044	g32DG2.5 as	aaacCGCCGTGAACAGTTCCTCGc
Cloning of CYBB sgRNAs		
BH137	sgCYBB54 s	caccgCAGCACTGGCACTGGCCCA
BH138	sgCYBB54 as	aaacTGGGCCAGTGCCAGTGCTGc
BH139	sgCYBB173 s	caccgGTACCTGGCTGTGACCTGT
BH140	sgCYBB173 as	aacACAGGTCACAGCCAGGTACc
BH141	sgCYBB45 s	caccgTTTACACAAGAAAACGTCT
BH142	sgCYBB45 as	aaacAGACGTTTTTCTTGTGTAAAc
BH143	sgCYBB124 s	caccgACATCTATTTAATGTGTAA
BH144	sgCYBB124 as	aaacTTACACATTAAATAGATGTc
DSA57	sgCYBB-Ex3 s	caccgTCAGCACTGGCACTGGCCA
DSA58	sgCYBB-Ex3 as	aaacTGGCCAGTGCCAGTGCTGAc
DSA60	sgCYBB-Ex6 s	caccgTACCTGGCTGTGACCCTGT
DSA61	sgCYBB-Ex6 as	aaacACAGGGTCACAGCCAGGTAc
Generation of EGFP mutant vectors		
SFHR037	SBmGW s	ATGGTGTCCAAGGGCGAGGAAGTTCACCGCGGCGTGGTGCCCATCCTG
SFHR038	SBmGW as	ATGGGCACCACGCCGCGGTGAACAGTTCCTCGCCCTTGGACACCATGGTT
SFHR046	SBmGW-K6 s	ATGGTGTCCAAGGGCGAGGAAGTTCACCGTGGTGCCCATCCTG
SFHR047	SBmGW-K6 as	ATGGGCACCACGTTGAACAGTTCCTCGCCCTTGGACACCATGGTT
SFHR048	SBmGW-K16 s	ATGGTGTCCAAGGGCGAGGAAGTTCACCGCGAGCGTGGTGCCCATCCTG

SFHR049	SBmGW-K16	as	ATGGGCACCACGCTCGCGGTGAACAGTTCTCGCCCTTGGACACCATGGTT
SFHR050	SBmGW-K18	s	ATGGTGTCCAAGGGCGAGGAAGTTCACCGCGGGCGTGGTGGCCATCCTG
SFHR051	SBmGW-K18	as	ATGGGCACCACGCCCCGCGGTGAACAGTTCTCGCCCTTGGACACCATGGTT
SFHR052	SBmGW-K29	s	ATGGTGTCCAAGGGCGAGGAAGTTCACCGCGGTGGCCATCCTG
SFHR053	SBmGW-K29	as	ATGGGCACCGCGGTGAACAGTTCTCGCCCTTGGACACCATGGTT
SFHR054	SBmGW-K33	s	ATGGTGTCCAAGGGCGAGGAAGTTCACCGCGCGGTGGCCATCCTG
SFHR055	SBmGW-K33	as	ATGGGCACCGCGCCGGTGAACAGTTCTCGCCCTTGGACACCATGGTT

Generation of CYBB mutant vectors

BH111	SB173CW	s	GGGCGGCCTGTACCTGGCTGTGACCTGTTGGCCGGCATCACCGGCGTG
BH112	SB173CW	as	CACGCCGGTGATGCCGGCCAACAGGTCACAGCCAGGTACAGGCCGCCC
BH113	SB54CW	s	GCTGCTGGGCTCAGCACTGGCACTGGCCCAGGGCCCCTGCCGCCTGCCTG AA
BH114	SB54CW	as	TTCAGGCAGGCGGCAGGGGCCCTGGGCCAGTGCCAGTGCTGAGCCCAGCA GC
BH121	SB124CW	s	GCGCCATCCACACCATCGCACATCTATTTAATGTGTAATGGTGCCTGAAC GCCCCGGGTG
BH122	SB124CW	as	CACCCGGGCGTTTACGCACCATTACACATTAAATAGATGTGCGATGGTGT GGATGGCGC
BH123	SB45CW	s	CGACATCCCCCCCCAAGTTCTTTTACACAAGAAAACGTCTTGGCAGCGCCC TGGCCCTGG
BH124	SB45CW	as	CCAGGGCCAGGGCGCTGCCAAGACGTTTCTTGTGTAAAAGAACTTGGGG GGGATGTCG

Shotgun cloning

AR36	Shotg HPRT	s	TCCCTTCATAGAGACAAGGAATG
AR163	Shotg HPRT	as	CCTGGGTTCTACCCCAGCACAG
AR102	Shotg EGFP	as	TCATGTGGTCGGGGTATCTG
SFHR20	Shotg EGFP	as	CTTGACAGCTCGTCCATGC
SFHR43	Shotg IRES	s	TCTGTAGCGACCCTTTGCAG
BH234	Shotg CYBB	as	AGGATCAGGCACAGGGTGAT

Surveyor assay

BH229	Svyr45/54	s	TGGCTCTCCTCAAGCGTATT
BH230	Svyr45/54	as	CCATCTTGTGGAAGGTCAGG
BH231	Svyr173	s	CACCTGTTCAACGTGGAGTG
BH232	Svyr173	as	CTCGCACAGGTACAGGAACA
BH233	Svyr124	s	AACGTGTTCTGTTCGTGTG
BH234	Svyr124	as	AGGATCAGGCACAGGGTGAT

2.2 Methods of molecular biology

2.2.1 Cultivation of bacteria

Liquid culture

E.coli (usually *DH5α*) bacteria were grown in Luria Broth (LB) medium supplemented with 100 µg/ml ampicillin (LB-Amp) at 37°C overnight (~16 h) and 220 rpm in a bacteria shaker (Roth).

Culture plate

Bacteria were plated on ampicillin containing LB-agar plats by using an inoculation spreader. Plates were incubated overnight at 37 °C in an incubator until bacteria colonies were visible (~16 h). Bacteria plates were stored up to 4 weeks at 4°C.

2.2.2 Cloning processes

Restriction

For restriction of DNA type II restriction endonucleases were used, which recognize specific sites of double stranded DNA sequences (usually 4 – 8 bp) and cleave in the sequence. Thereby either blunt or sticky ends are produced. DNA digestions were performed using commercially available endonucleases from New England Biolabs (NEB) according to the manufacturer's instructions. For a preparative restriction 1 – 5 µg and for an analytical restriction 0.2 – 1 µg DNA was used.

Standard digestion:

0.2 to 5 µg	DNA for analytical or preparative purpose
0.1 µl volumes	10x reaction buffer (corresponding NEB buffer 1-4)
10 U per µg DNA	Restriction enzyme
add till 0.05 µg/µl DNA	ddH ₂ O

The reaction mixture was incubated for 1 – 2 h at 37°C or higher (enzyme dependent). In case of a double digest the optimal buffer for the double digest was chosen according to the manufacturer's instructions.

Dephosphorylation of 5'-ends of DNA

The antarctic phosphatase catalyzes the removal of 5' phosphate groups from DNA and RNA. Dephosphorylation was used to minimize self-ligation of the vector background in cloning reactions. Since phosphatase-treated fragments lack the 5' phosphoryl terminus required by ligases, they cannot self-ligate. The enzyme can be added directly into the digestion reaction after DNA digestion.

Standard dephosphorylation:

50 µl	Restriction mix
6 µl	Antarctic Phosphatase-Buffer (10x)
2 µl	Antarctic Phosphatase
add 60 µl	ddH ₂ O

The reaction was incubated for 30 min at 37°C.

Purification and ligation of vectors and inserts

Before ligation, the vectors and the inserts were purified with the "DNA Clean & Concentrator"-kit (Zymo Research) according to the manufacturer's instructions. For ligation, the T4 DNA ligase was used to catalyze the formation of phosphodiester bonds between the fragments under consumption of ATP. The ligation was applied for fragments exhibiting complementary overhangs (sticky-end ligation) or blunt ends (blunt-end ligation). As control sample a reaction which only contains plasmid DNA was used.

Standard ligation:

0.1 – 1.0 µg	total DNA (3:1 to 5:1 Insert:Vector ratio)
2 µl	10x reaction buffer
1 µl	T4 DNA ligase
add 20 µl	ddH ₂ O

The reaction mixture was incubated at RT for a minimum of 2 h before transformation into competent bacteria.

TA-Cloning

This easy and fast cloning technique relies on the ligation of the overhang of adenine and thymine at the end of different DNA fragment. Shot gun cloning of PCR amplification products was performed using the pGEM-T vector system (Promega, Mannheim, Germany) according to the manufacturer's instructions. pGEM-T vector is a linearized vector with a single 3'-terminal thymidine at both ends. These compatible overhangs for PCR products generated by certain thermostable polymerases (e.g. Taq-Polymerase) greatly improve the efficiency of ligation. Positive clones were sequenced with a T7 primer.

Standard reaction

1 – 4 µl	PCR-reaction
1 µl	10x reaction buffer
1 µl	pGEM-T vector
add 10 µl	ddH ₂ O

The reaction was incubated at RT for 1 h. The reaction was subsequently used for transformation in competent bacteria.

Golden Gate Assembly

Golden Gate assembly exploits the ability of type IIS restriction endonucleases to cleave DNA outside of the recognition sequence. Combinations of overhang sequences can be assembled using T4 DNA ligase. This one step cloning method combines restriction and ligation reaction in one tube (Engler et al. 2008¹³⁹).

Annealing of the oligonucleotides

50 pmol of each complementary oligonucleotides were added to 40 µl ddH₂O and incubated at 98°C for 5 min. Annealing was achieved by cooling down the oligonucleotide mixture at RT for 20 min.

Standard oligo annealing

5 µl	Oligo 1 (100 µM)
5 µl	Oligo 2 (100 µM)
40 µl	ddH ₂ O

Golden Gate reaction

150 ng	Vector (LCv2 and derivatives)
5 µl	Annealed Oligos
2 µl	10x ligase Buffer
1 µl	T4 DNA ligase
1 µl	BsmBI restriction enzyme
add 20 µl	ddH ₂ O

Program

37°C	5 min	} 10 cycles
16°C	10 min	
37°C	15 min	1 cycle
80°C	5 min	1 cycle
4°C	Hold	

2.2.3 Transformation of competent bacteria

Transformation of *E. coli* (usually *DH5α*) is the method to amplify plasmid DNA through cellular replication. Most of the plasmids are “high copy” plasmids which are amplified to multiple copies within the bacteria independent of cellular replication resulting in high yields of plasmid DNA.

Chemically competent DH5α cells were thawed on ice and approximately 10 ng DNA or 20 µl ligation reaction was added and incubated on ice for 20 min. A heat shock was performed by incubation in a water bath at 42°C for 1 min. After additional on ice for 5 min, cells were plated on agar plates containing ampicillin (100 µg/ml) and incubated at 37°C overnight.

2.2.4 Plasmid preparation

Miniprep

For analytical scale purification of low amounts of DNA (Miniprep), 2 ml LB-Amp medium were inoculated with single bacteria clone in a 10 ml tube (Falcon) and grown at 37°C overnight. The bacteria were pelleted at RT and 13000 rpm for 5 min in a benchtop centrifuge and the supernatant was discarded. The bacteria pellet was processed with buffers from the Maxiprep-Kit (Nucleobond PC500 Maxi Kit). In brief, the pellet was resuspended in 100 µl S1 buffer, lysed 200 µl S2 buffer for 5 min at RT and neutralized using 150 µl S3 buffer. Debris was pelleted at 13000 rpm for 5 min. The supernatant was transferred to a fresh tube and the DNA precipitated by addition of 1 ml of 100% ethanol. The DNA was pelleted at 13000 rpm for 5min, washed once with 150 µl of 70% ethanol (13000 rpm for 5min). The dried pellet was resuspended in 50 µl ddH₂O.

Maxiprep

For extraction of larger amounts of DNA (Maxiprep) 250 ml LB Amp medium were inoculated and cultivated at 37°C overnight in a 1500 ml shake flask. Preparative scale plasmid preparation from transformed bacteria was performed using the Nucleobond PC500 Maxi Kit (Macherey-Nagel) according to the manufacturer's instructions. This kit uses an anion-exchange tip where plasmid DNA selectively binds under low-salt and pH conditions. RNA, proteins, metabolites, and other low-molecular-weight impurities are removed by a medium-salt wash, and pure plasmid DNA is eluted in high-salt buffer. The DNA is concentrated and desalted by isopropanol precipitation and collected by centrifugation. The DNA pellet was dried and dissolved in ddH₂O to a concentration of 1 µg/µl. The concentration of the isolated plasmid DNA was determined photometrically at absorption A₂₆₀ by NanoDrop (Thermo Fisher Scientific).

2.2.5 Agarose gel electrophoresis

Agarose gel electrophoresis allows the separation of DNA molecules by their size. 1 - 2% agarose gels were used, the required amount of agarose powder was mixed with 1x TAE buffer and dissolved by heating the emulsion in a microwave oven. 0.1 µg/µl ethidium bromide or 1 µl HDGreen Plus per 20 ml agarose gel was added to the dissolved agarose and the gel was casted into a tray. Ethidium bromide / HDGreen

Plus intercalates into DNA and visualizes DNA under UV light. Combs with the appropriate number and size of teeth were used to make the loadings slots. After polymerization, the gel was covered in 1x TAE buffer, samples were mixed with 10x loading buffer and loaded onto the gel. 1 kb Plus DNA Ladder was used as marker. The DNA fragments were separated by applying 120 V for 30 – 60 min in an electrophoresis chamber. The DNA fragments were visualized on a transilluminator and documented or extracted from the gel (see below).

2.2.6 Gel extraction of DNA fragments

To isolate an appropriate DNA fragment from an agarose gel, the DNA band of interest was cut out of the gel under UV light and transferred into a 2 ml micro tube. The DNA purification from an agarose gels was performed using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research) according to the manufacturer's instructions. This kit is based on binding of DNA to silica based matrixes in the presence of a high concentration of chaotropic salt. After washing the agarose to remove impurities the pure DNA was eluted using 50 µl of ddH₂O.

2.2.7 Polymerase chain reaction (PCR)

The polymerase chain reaction (PCR) is a method for oligonucleotide primer directed enzymatic amplification of a specific DNA sequence of interest. The DNA polymerase amplifies a template DNA bordered on the forward and reverse primer in an exponential manner. The polymerase need appropriate buffer, oligonucleotides (primers), deoxynucleotides (dNTPs) and cycling conditions. The amplified fragment can be used for further molecular biological methods.

Standard PCR-reaction

10 µl	10x Puffer
2 µl	10 mM dNTP-Mix
2.5 µl	Forward Primer (10 µM)
2.5 µl	Reverse Primer (10 µM)
50 – 500 ng	Template DNA
0.5 µl	Q5 High-Fidelity DNA Polymerase
add 50 µl	ddH ₂ O

Standard PCR-protocol

98°C	2 min	Initial Denaturation	
98°C	20 sec	Denaturation	} 25 – 35 Cycles
50 – 72°C	20 sec	Primer Hybridization	
72°C	30 sec/kb	Elongation	
72°C	2 min	Final Extension	
4°C	∞	Hold	

The optimal primer hybridization temperature was determined by OligoAnalyzer from NEB (<http://tmcalculator.neb.com/#!/>) and in some cases by a temperature gradient. Amplificated products were subsequently analyzed by agarose gel electrophoresis (2.2.6).

2.2.8 PCR product extraction

The PCR-product was purified using the DNA Clean & Concentrator™ -5 (Zymo Research) according to the manufacturer's instructions. This kit is based on binding of DNA to silica based matrix in the presence of a high concentration of chaotropic salt. After washing off the impurities the pure DNA was eluted using 50 µl of ddH₂O.

2.2.9 Nucleic acid sequencing

Nucleic acid sequencing was performed by SeqLab Company (Göttingen, Germany). DNA samples containing approximately 1 µg plasmid or 0.1 µg PCR-product and 30 pmol sequencing primer in 15 µl total volume were prepared and sent to the sequencing service.

2.2.10 Isolation of genomic DNA from cells

Phenol/Chloroform/Isoamylalkohol (25:24:1) was used for gDNA extraction. Up to 5x10⁶ cells were centrifuged 5 min at 1200 rpm at RT and the supernatant was discarded. The cell pellet was resuspended in 400 µl SDS cell-lysis buffer, containing 150 µg/ml proteinase K. The cells were incubated at 60°C overnight. Afterwards, 300 µl phenol/chloroform/isoamylalkohol were added and the mixture was vortexed and then centrifuged (RT, 13000 rpm, 5 min). The aqueous phase containing DNA (upper) was

transferred to a new tube and mixed with 300 µl chloroform. Vortexing and centrifugation was repeated. The upper phase was transferred to a new tube and mixed with 1 ml 100% ethanol (vortexed and centrifuged) for precipitation of the DNA. The pellet was washed with 500 µl of 70% ethanol (13000 rpm for 5min at RT). The genomic DNA pellet was dried and dissolved in ddH₂O to a concentration of 200 ng/µl.

2.2.11 Surveyor Assay

For the analysis of CYBB indels generated by engineered nucleases, the enzyme mismatch cleavage method was used. This method uses enzymes, such as T7E1 or Surveyor, that cleave heteroduplex DNA at mismatches and extrahelical loops formed by single or multiple nucleotides. In this study, Surveyor® Mutation Detection Kit-S100 (IDT, Leuven, Belgium) was used. RNG targeted CYBB sites were PCR amplified to obtain 500 nt sized products and PCR product were purified using the DNA Clean & Concentrator™ -5 Kit (Zymo Research). For the analysis of on-target mutation rates, 400 ng of purified PCR product was subjected to the Surveyor assay according to the manufacturer's instructions. PCR products were mixed with 2 µl 10× Taq DNA Polymerase PCR buffer and water was added to a final volume of 20 µl. To enable heteroduplex formation the DNA was reannealed as follows: 95°C for 10 min, 95°C to 85°C ramping at -2°C/s, 85°C to 25°C at -0.25°C/s, 25°C for 1 min and hold at 4°C. After re-annealing, products were treated with Surveyor nuclease and Surveyor enhancer S following the manufacturer's instructions and analyzed on 2% Agarose gel (see 2.2.5). Indel percentages derived from Image J plots¹⁴⁰ were calculated by following formula¹⁰⁵:

$$100 * \left(1 - \left(1 - \frac{b + c}{a + b + c}\right)^{0.5}\right)$$

Where a is the integrated intensity of the undigested PCR product, and b and c the integrated intensities of each cleavage product.

2.2.12 Southern Blot

Southern blot is a technique to detect specific DNA sequences in genomic DNA samples. It combines the transfer of electrophoretically separated DNA fragments to a nylon membrane and subsequent detection of specific DNA sequences by hybridization with radioactively labeled DNA probes.

DNA isolation was performed as described in chapter 2.2.10. 15µg of genomic DNA were digested overnight using NcoI restriction enzyme and fragments were resolved on a 0.8% agarose gel alongside a DNA size standard. The gel was stained for 30min in 1x TAE + EtBr and a picture was taken for later size identification of visualized DNA fragments. The gel was destained in depurination solution (0.25 M HCl) for 15 min and rinsed in water. Thereafter, the gel was neutralized in alkaline transfer buffer (0.5 M NaOH, 1.5 M NaCl) for 30 min and an alkaline transfer was set up for the transfer of the DNA fragments onto a Nytran SPC membrane (Whatman, Maidstone, UK). Therefore, a pack of disposable paper towels was set up as a suction reservoir. On top of this pack two Whatman layers were placed with the same size as the gel on which the membrane was mounted. The gel was directly placed onto the membrane avoiding air bubbles. The whole transfer set-up was covered with plastic wrap, the gel was weighted with a plastic plate and a bottle with approx. 500ml of water. Transfer was performed over 48h. After transfer, the blot set-up was broken down carefully, discarding all paper towels, Whatman paper and the dried gel. The membrane was neutralized for 15min in 0.5M Tris/HCl pH 7.5, 1.5 M NaCl and equilibrated for hybridization in 2xSSC (saline sodium citrate buffer). Radioactive labelling of the BFP probe was carried out using the RediPrime II DNA Labeling System (Amersham) according to the manufacturer's protocol. Hybridization was performed using ExpressHyb Hybridization Solution (Clontech) according to the manufacturer's protocol. The membrane was washed twice at 37°C with 2xSSC, 0.1% SDS and twice with 0.5xSSC, 0.1% SDS. Signals were detected using overnight exposure of Kodak BioMax MS films at -80°C using amplifier screens.

2.2.13 Western Blot

Western blot is a technique to detect specific proteins from a complex mixture of proteins extracted from cells. The technique is based on four elements: separation protein by size, protein transfer to a membrane and target protein with specific antibodies and detection of the specific signals.

Cell lysis to extract protein:

Cells were harvested by centrifugation at RT and 1200 rpm for 5 min and the supernatant was discarded. The cell pellet was washed once with 1xPBS and the pellet was lysed in SDS-lysis buffer (75 – 250 µl) supplemented with Protease Inhibitor Cocktail (Roche, Mannheim, Germany) on ice for 30 min. The cells were then sonicated twice for 30 sec, (5x, 45% power) and transfer to ice. The cool lysates mixture was centrifuged at 4°C and 15000 rpm for 3 min. The supernatant was transferred to a new 1.5 ml tube.

Protein concentration was determined using the Lowry protein assay. To measure protein concentration, 1 – 5 µl sample or protein standard was mixed with 200 µl Lowry solution A, filled up to 1 ml with Lowry solution B, mixed by vortexing and samples were measured at the absorbance 595 nm. The samples with equal protein concentrations were resuspended in 4x loading buffer containing 20% beta-mercaptoethanol, boiled at 95°C for 5 min and stored at 20°C.

SDS polyacrylamide gel electrophoresis (SDS-PAGE) and transfer:

The proteins were separated by SDS-PAGE. Equal volumes of protein lysates were loaded onto a polyacrylamide gel (stacking gel: 4% acrylamide, separating gel: 10% acrylamide) together with 1 µl protein standard. The gel was run in MOPS buffer, starting with 80 V for 10 min and then increased voltage to 160 V for 1h

Component	4% Stacking gel	10% Separating gel
Acrylamide 40%	687 μ l	1500 μ l
Bis-Tris Buffer	550 μ l	1730 μ l
ddH ₂ O	2720 μ l	2800 μ l
APS 10%	40.3 μ l	40 μ l
TEMED	4 μ l	11.2 μ l

After separation by SDS-PAGE, western blot were blotted via electrophoresis onto a nitrocellulose membrane using the wet blot method. The membrane and 2 filter papers were cut to gel size. Membrane, filter paper and blotting pads were soaked in pre-cooled (4°C) transfer buffer. Blotting pads were placed on cathode, on topped with a filter paper, the gel was removed from the electrophoresis apparatus and placed on top, before membrane, another filter paper and more blotting pad were placed. After each layer roll out any air bubbles. The transfer was carried out in transfer buffer at 35 V and 500 mA for 2 h.

Immunostaining of blotted proteins and detection:

To get a first impression of the quality of the transfer and the equality of protein loading, an unspecific protein staining with ponceau-red was performed. The membrane was blocked with 5% milk in TBS-T at RT for 1h, before incubated with a target protein specific primary antibody on a shaker at 4°C overnight. On the next day the membrane was washed 3 times for 5 min with TBS-T and the secondary antibody was added at RT for 1 h. Before detection, the membrane was washed again 3 times with TBS-T. The ECL-substrate for horseradish peroxidase was applied to the membrane according to manufacturer`s instructions and detection was performed with the Odyssey FC device.

Mouse monoclonal antibody against GFP was purchased from Roche and rabbit monoclonal antibody against GAPDH was purchased from Cell Signaling (Frankfurt am Main, Germany). Monoclonal anti-human CYBB antibody (moAB48) was obtained from LifeSpan BioSciences (Seattle, WA, USA).

2.3 Cell culture and virological methods

2.3.1 Cultivation of cell lines

Cell lines were cultivated in the appropriate medium in an incubator at 37°C, 5.0% CO₂ in a saturated water atmosphere and were passaged twice a week. Hematopoietic (Suspension) cells were grown in RPMI-1640 medium supplemented with 10% (v/v) heat-inactivated fetal calf serum (FCS), 2 mM glutamine, 100 U/ml penicillin and 100 µg/ml streptomycin. Hematopoietic 32D cell cultures were supplemented further with 10% (v/v) WEHI-3B conditioned medium. WEHI-3B cells were grown in IMDM supplemented with 10% (v/v) heat-inactivated fetal calf serum (FCS), 2 mM glutamine, 100 U/ml penicillin and 100 µg/ml streptomycin, 0,1% beta-mercaptoethanol. For splitting the cells were washed with 1xPBS by centrifugation (at 1200 rpm, 5 min) and an appropriate cell number was resuspended in fresh media.

The adherent HEK293T cells were grown in DMEM medium supplemented with 10% FCS (v/v), 2 mM glutamine, 100 U/ml penicillin and 100 µg/ml streptomycin. The cells were grown until a confluence of 80 – 90% was reached and then split by washing once with PBS and detaching with Trypsin-EDTA before an appropriate fraction of the resulting suspension was seeded into a new culture flask with fresh medium.

2.3.2 Cell counting and determination of cell viability

Cell number and viability were controlled by trypan-blue exclusion. To count the cells 10 – 50 µl cell suspension was mixed with an equal volume of trypan-blue solution. Trypan-blue selectively stains dead cells as it is not able to pass through intact cellular membranes. 10 µl of the mixture was applied to a counting chamber and only unstained cells were counted. Cell number were calculated according to this formula:

$$Cell\ number/ml = \frac{counted\ cells}{number\ of\ counted\ squares} * dilution\ factor * 10^4$$

2.3.3 Make WEHI-3B conditioned media

WEHI-3B cells were expanded in 150-cm² flasks to a density of 80% to 90% confluence. The cells were collected, centrifuge (at 1200 rpm and 5 min) and washed once with 1xPBS (at 1200 rpm, 5 min). The cell pellet divided to four 150-cm² tissue culture flasks, each with a total volume of 150 ml of growth medium. During the expansion stage, the pH indicator in the conditioned medium might have become yellowish in color. 90% conditioned medium was collected every 2 – 3 days and filtered through a 0.22 µm filter. Fresh medium was added to the remaining cells. Filtered conditioned medium was collected in sterile 50-ml conical tubes in 50-ml aliquots to and stored at –20°C.

2.3.4 Freezing and thawing of cultured cell

Freezing

To freeze cells, adherent cells were first washed once with PBS, detached through trypsinising, resuspended in the appropriate medium and then like suspension cells, centrifuged (1200 rpm, 5 min) to pellet the cells. The cell pellet was washed with 1xPBS, resuspended in freezing medium (90% FCS, 10% DMSO) to reach 5x10⁶ cells per 1 ml, transferred to cryotubes in 1 ml aliquots and frozen in a freezing container at -80°C. After 3 days the cells were transferred into liquid nitrogen.

Thawing

Cryotubes were quickly placed at 37°C until the suspension was almost thawed. Immediately, the cells were transferred into a 50 ml falcon tube with 24 ml prewarmed medium. The cells were centrifuged at 1.200 rpm for 5min to remove the cytotoxic DMSO. Lastly the cells were resuspended in fresh medium and seeded into appropriate culture flasks. 24 h later the medium was exchanged to fresh medium.

2.3.5 Nucleofection

Nucleofection is an electroporation-based transfection technique. Electroporation is a biophysical method in which the cell membrane becomes permeable by applying pulsed electric fields across a cell. This allows chemicals, protein or DNA to be introduced into the cells. The nucleofection technology is carried out with cell type

specific electrical parameters, generated by a device called Nucleofector and with cell-type specific solutions.

For nucleofection of 32D cells, 1×10^6 cells were washed twice in 1xPBS, resuspended in nucleofection buffer and nucleofected with 1 µg donor plasmid, 0.5 µg of each Hprt-TALEN or 0.4 µg hspCas9 and 1.6 µg of sgRNA expression vector (Amaxa nucleofector kit V with E-32 program, Amaxa Biosystems, Cologne, Germany) according to the manufacturer's protocol. The cells were selected in growth medium containing either 2 µg/ml 6-thioguanin (6-TG) (Sigma-Aldrich, Taufkirchen, Germany) or 1 x hypoxanthine-aminopterin-thymidine (HAT) (Life technologies, Darmstadt, Germany).

2.3.6 Lentiviral vector production

Lentiviral supernatants were produced by polyethylenimine (PEI) based transient co-transfection of HEK-293T cells. Briefly, transfer vector, lentiviral *gag/pol* helper plasmids for integration (pCMV8.91) or integrase-defective (pCMV8.74) variants and envelope plasmid encoding the glycoprotein of vesicular stomatitis virus (VSV-G) (pMD2.G, Addgene #12259) were transfected at a molar ratio of 3:1:1 by standard PEI transfection. 24 h prior to transfection, HEK 293T cells were seeded at 1.2×10^7 per 14.5 cm plate in 16.5 ml DMEM with 10% FCS (v/v), 2 mM glutamine, 100 U/ml penicillin and 100 µg/ml streptomycin to have a confluence 75-85% at the time of transfection. At the day of transfection:

Prepare DNA solution in a 14 ml tube for every plate:

14 µg	psPAX2 (gag, pol)
8 µg	pMD2.G (vsvg)
16 µg	Transfer DNA
1.75 ml	DMEM w/o supplements

Prepare PEI solution in a 14 ml tube for every plate

171 µl	PEI (1 µg/µl PEI → 4.5 µl per 1 µg DNA)
1.75 ml	DMEM w/o supplements

After 12 – 16 h the medium was replaced with 11 ml fresh medium. A maximum of 33 ml per viral supernatant could be used which was routinely collected from three 14.5cm dishes. After another 48 h viral supernatants were harvested, the supernatants of the same viral vector were pooled, sterile filtered (0.45 µm pore-size PVDF-membrane filter; Millipore, Schwalbach, Germany) and concentrated (100-fold) by ultracentrifugation. The virus was underplayed with 5 ml of a sucrose (20% v/v sucrose in PBS) and centrifuged (50,000 g, 2:20 h, 4°C). Pelleted viral particles were resuspended in 330 µl StemSpan SFEM serum-free medium (StemCell Technologies, Grenoble, France) without any supplements at 4°C overnight and stored at –80°C in 50 µl aliquots.

2.3.7 Titration of vector particles

Several different methods for titration of viral supernatants exist. In the present study two different methods were used.

For integrated lentivirus, which expresses a virus encoded surface transgene in target cells, the functional titer was determined by serial dilutions of viral supernatant transduced of XCGD-PLB985 cells followed by flow cytometry 3 days post-transduction. The number of cells expressing the vector-encoded transgene (CYBB or TagBFP/EGFP) directly correlates with the number of functional vector particles in the supernatant and can therefore be calculated. For the titration of virus, 2×10^5 XCGD-PLB985 cells were seeded into a single well of a 24well plate. The lentiviral suspension was serially diluted in 1:10 steps with medium supplemented with 8 µg/ml polybrene as a chemical transduction enhancer and applied to the wells to a final volume of 500µl medium. After incubation overnight, media was removed and 1 ml fresh medium was added. Cells were grown for 3 days before analysis of the percentage of positive cells by flow cytometry. For CYBB, a FITC labelled Anti-Flavocytochrome b 558 monoclonal antibody which recognizes an external epitope of CYBB (7D5-FITC¹⁴¹) was used. For vector titer calculation, dilutions, that contained about 5 - 20% of transduced cells, was chosen.

Calculation of virus titer:

$$\text{Titer} \left[\frac{\text{TU}}{\text{ml}} \right] = \left(\frac{\text{number of seeded cells} * \text{dilution factor} * \text{positive cells [\%]}}{100 [\%]} \right) / 0.5 \text{ ml}$$

To determine IDLV titers, p24 viral coat protein concentrations were estimated in the viral supernatants using the p24 ELISA kit from INNOTEST (Fujirebio, Hannover, Germany) according to the manufacturer's instructions. To determine the number of particles corresponding to every picogram of p24 antigen, we used a conversion factor of 6.12×10^3 particles/pg derived from the flow cytometric analysis of PLB cells 48 h after SBmGW-IDL V transduction.

2.3.8 Flow cytometry and cell sorting

Flow cytometry, also referred as Fluorescence Activated Cell Sorting (FACS), is a biophysical technique for biomarker detection. An appropriate amount of cells ($0.1-1 \times 10^6$) from a cell culture were transferred to 5 ml polystyrol-FACS-tubes. The tubes were filled to the top with PBS and centrifuged at 1200 rpm for 5 min. The supernatant was removed, the cells were resuspended in 200 µl PBS and ready for measurements or staining. For cell surface antigen staining, cells were incubated with conjugated antibodies (1 µl of each antibody) in dark, at RT for 20 – 30 min. Then, one wash step was performed with PBS to remove unbound antibodies. For the FACS measurement the cells were kept in 200 µl PBS. Data acquisition was performed with a *BD LSRFortessa* flow cytometer (BD Biosciences, Heidelberg, Germany). Data was analyzed with BD FACSDiva software (BD Biosciences, Heidelberg, Germany) or flowing software 2.5.1. Cell sorting was performed in a BD FACSAria III flow cytometer (BD Biosciences, Heidelberg, Germany).

2.3.9 Dihydrorhodamine 123 (DHR) reduction assay

The Dihydrorhodamine 123 (DHR) -assay is used to detect superoxide molecules generated by the NADPH-oxidase complex. It is a functional assay of the NADPH oxidase and can be used to investigate reconstitution of CYBB repair. The assay relies on the conversion of non-fluorescent DHR123 to highly fluorescent Rhodamine123 preferentially but not exclusively by superoxide. In addition to superoxide, its

derivatives, such as hydroxyl anions or peroxide can also convert DHR, albeit to a lesser extent. Activation of NADPH-oxidase complex in target cells is triggered by the addition of the unspecific PKC-activator PMA (Phorbol 12-myristate 13-acetate).

DHR assays were performed as described by Brendel et al 2014¹⁴². For granulocytic differentiation, PLB-985 and its derivatives were plated in RPMI 1640 supplemented with 2.5% heat inactivated FCS, 2 mM glutamine, 100 U/ml penicillin and 100 µg/ml streptomycin and 1.25% dimethyl sulfoxide (DMSO) at a concentration of 2×10^5 cells/ml for at least 7 days. Flow cytometry was performed to ensure the differentiation of labelled cells with anti-CD11b⁺ antibody. For estimating ROS production, the differentiated cells were suspended in 1 ml pre-warmed HBSS (Life technologies, Darmstadt, Germany) supplemented with 7.5 mmol/l D-Glucose, 0.5% bovine serum albumin (BSA; Sigma-Aldrich) 2000 U/ml catalase (Sigma-Aldrich) and 5 µg/ml DHR123 (Sigma-Aldrich). Following incubation at 37°C for 10 min, cells were exposed to 0.1 µmol/l PMA (Sigma-Aldrich) for 15 min and placed on ice. Rhodamin123 fluorescence was measured in a flow cytometer within the next 30 min.

2.3.10 Statistical analysis

For statistical comparisons between groups, Student's t-test or one-way analysis of variance (ANOVA) with Bonferoni post-hoc test were used as appropriate in conjunction with GraphPad Prism 5 software.

3. Results

3.1 Genome editing by homology-directed repair (HDR) versus non-homologous end joining (NHEJ)

In an initial project aimed at testing the efficacy of genome editing in hematopoietic cells, Transcription activator like effectors nucleases (TALENs) were used. Two TALENs were designed directed against 16 nucleotide (nt) stretches on both strands of the 3rd murine *Hprt* exon encoding for the active centrum of hypoxanthine-guanine phosphoribosyltransferase (Hprt) enzyme. The Hprt-TALENs were separated from each other by a 15 nt spacer (Figure 10A). 32D cells (murine hematopoietic, immature, myeloid cell line) were nucleofected with Hprt-TALENs and *EGFP* encoding donor constructs. The nucleofected 32D cells were selected against Hprt expression in 6-thioguanine (6-TG) containing medium. Less than 5% of the 6-TG resistant cells expressed EGFP from the knock-in donor construct (Figure 10B), suggesting that in the majority of cells *Hprt* inactivation was caused by indels associated with NHEJ-directed DSB-repair.

To confirm this, individual 6-TG resistant clones were isolated by limiting dilution. The mutant *Hprt* alleles in clones were characterized by genomic PCR and sequencing. Consistent with NHEJ, the *Hprt* open reading frame (ORF) was disrupted by indels of different sizes (Figure 10C).

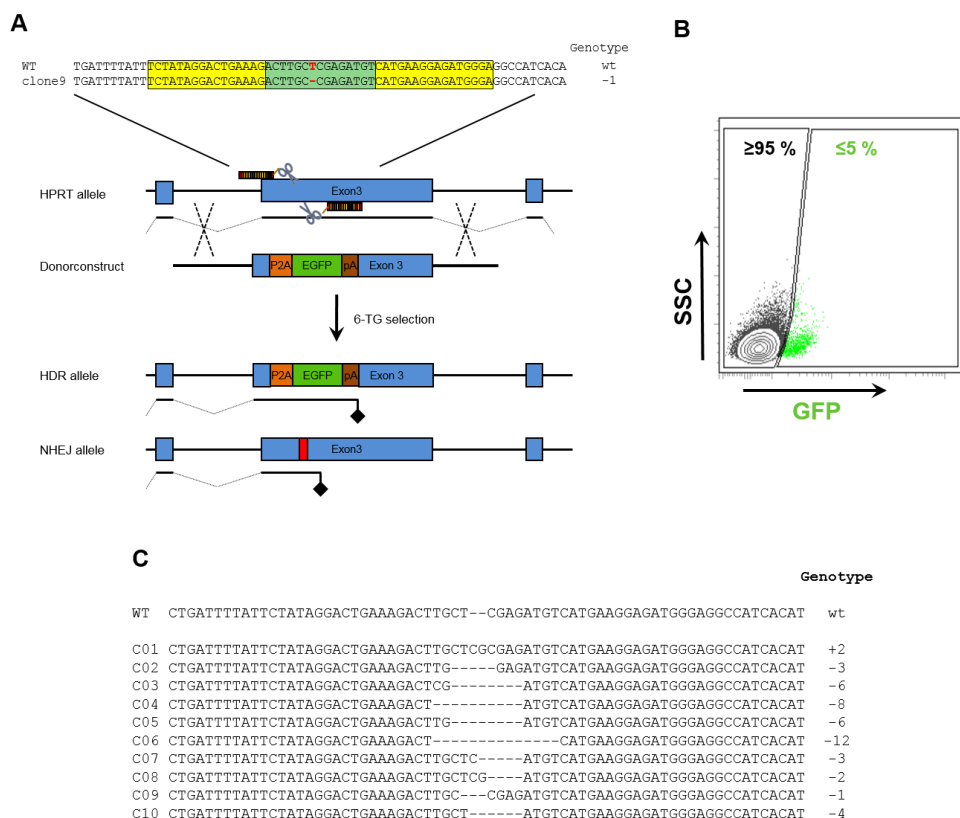


Figure 10: TALEN-assisted *Hprt* gene editing

A. Schematic illustration of exon 3 of *Hprt* gene with TALEN target sequence (highlighted in yellow) and EGFP encoding *Hprt*-targeting (donor) construct. **B.** Flow cytometry of 6-TG resistant 32D. Note that less than 5% of the 6-TG cells underwent HDR, suggesting a NHEJ/HDR ratio of 19/1. **C.** Indel sequences recovered from 6-TG resistant subclones.

3.2 Reconstitution of a point mutated *Hprt* gene by genome editing in murine hematopoietic (32D) cells

As shown in Figure 10C, modifications induced by NHEJ typically consist small deletions or insertions (indels) of variable length of which theoretically about one-third should restore the ORF disrupted by a disease mutation (Figure 11). This could lead to many ORF reconstitutions, of which some, depending on the position and type of the original mutation, should completely or partially recover protein function.

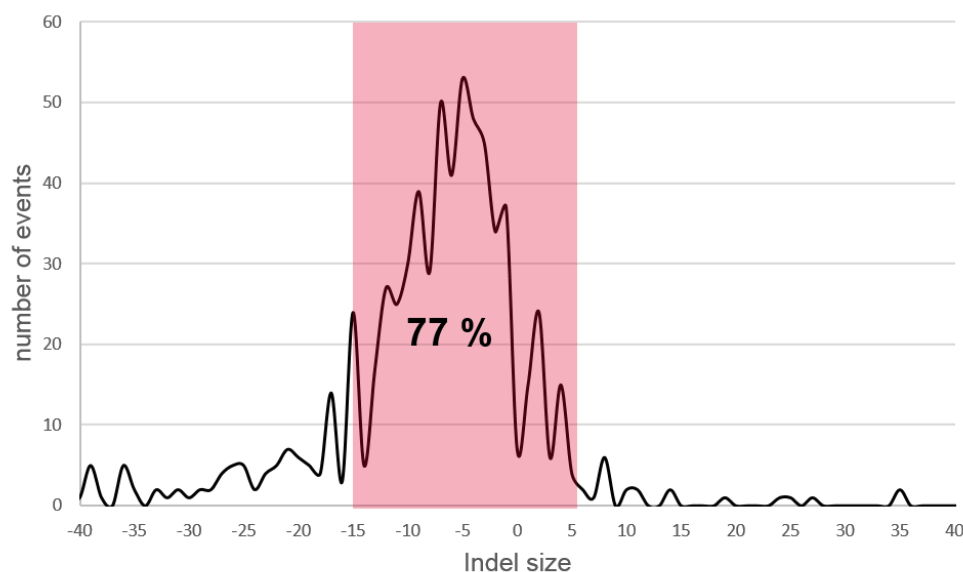


Figure 11: Analysis of INDEL distribution

Plot depicting frequency of ~750 indels evoked by designer endonucleases from different publications. If one postulates that one third of indels between -15 and +5 nt could reconstitute the reading frame of a mutated gene, in approximately 25% of all targeted cells expression of the gene will be restored. Note: 92% of all indels are between -40 - +40 nt.

To test this prediction, the 32D-C9 subclone (Figure 10C), harbouring a frameshift mutation in the active center of *Hprt*, was nucleofected with the same *Hprt*-TALENs. After positive selection in hypoxanthine-aminopterin-thymidine (HAT) medium, HAT-resistant cells were obtained, suggesting functional *Hprt* reconstitution. Genomic PCR and sequencing proved that the HAT resistant cells originated from two independent clones: one with a synonymous substitution restoring the Ala⁵⁰ codon and the other with Arg⁵¹ to Gly codon (Figure 12).

	E R L A R D V M K E M	
WT	GAAAGACTTGCTCGAGATGTCATGAAGGAGATG	
	E R L A E M S Stop	
Mut	GAAAGACTTGCCGAGATGTCATGAAGGAGATGG	
	E R L A R D V M K E M	
Repair-1	GAAAGACTTGCCGAGATGTCATGAAGGAGATG	→ Gene repair
	E R L A G D V M K E M	
Repair-2	GAAAGACTTGCCGAGATGTCATGAAGGAGATG	→ Kelley-Seegmiller-Syndrome

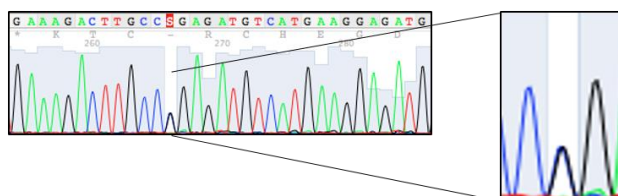


Figure 12: TALEN-assisted random *Hprt* gene repair

Indel sequences recovered from HAT resistant 32D-C9 subclone after nucleofection with Hprt-TALENs and followed by selection in HAT medium.

3.3 NHEJ-mediated reconstitution of a mutated *EGFP* reporter gene in 32D cells

To determine the efficiency of gene repair by NHEJ, 32D cells that underwent HDR after Hprt-TALEN and EGFP donor treatment (32D-EGFP+) were first enriched by fluorescence activated cell sorting (FACS) (Figure 10B & Figure 13A).

Using RNA-guided CRISPR/Cas9 endonuclease (RGN) technology, EGFP inactivating (*knock-out*) mutations were induced in 32D-GFP+ cells using two different single guide RNAs (sgRNAs) directed against the N-terminus of EGFP (Figure 13B). Expression of sgRNA1 (G1) or sgRNA2 (G2) plus spCas9 in 32D-GFP+ cells inactivated EGFP in 17.2% (32DG1) and 22.3% (32DG2) cells, respectively (Figure 13C).

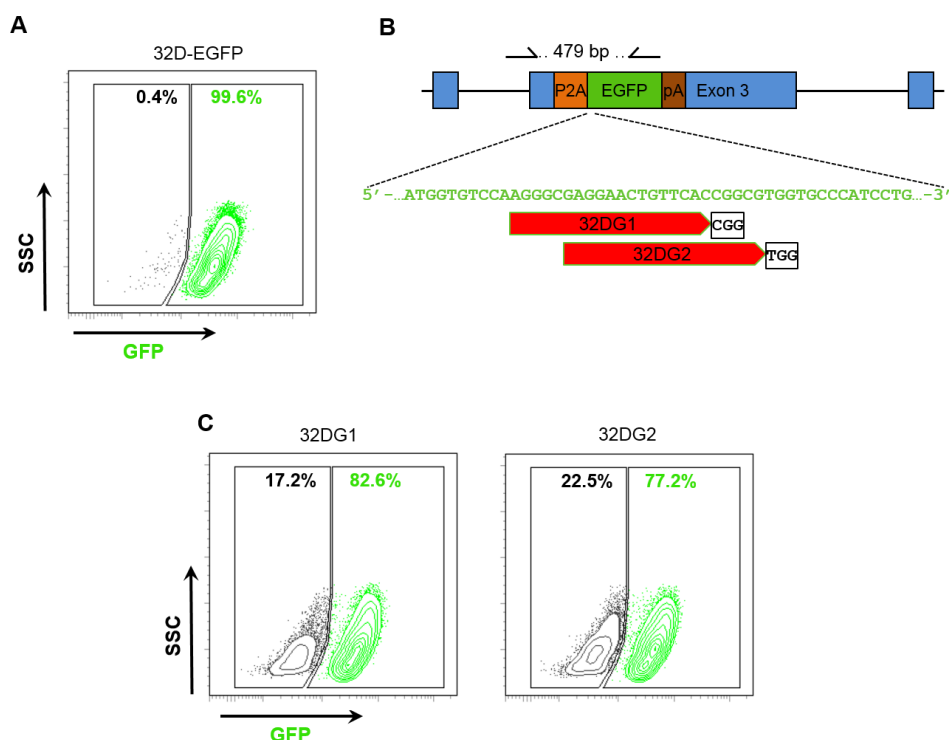


Figure 13: EGFP mutagenesis in 32D-GFP⁺ cells

A. FACS enrichment of 32D cells expressing EGFP from the *Hprt* locus. **B.** sgRNAs directed against EGFP. **C.** Frequency of EGFP inactivation in 32D-GFP⁺ cells after conucleofecting the sgRNAs and Cas9 expression constructs.

Next, ten single cell clones from sorted EGFP negative 32DG1 and 32DG2 cells were isolated by limiting dilution and their genomic DNA subjected to PCR using Hprt and EGFP specific primers (Figure 13B). As shown in Figure 14A, seven of the 32DG1- and six of the 32DG2-DNAs yielded amplification products of the expected size. The other clones either lacked the entire target sequence (32DG1.5 and 32DG2.9) or the generated product was significantly smaller than the expected size (32G1.10, 32DG2.2 and 32DG2.6), suggesting larger deletions (Figure 14A). To characterize the EGFP inactivating indels, PCR products of the expected size were sequenced (Figure 14B).

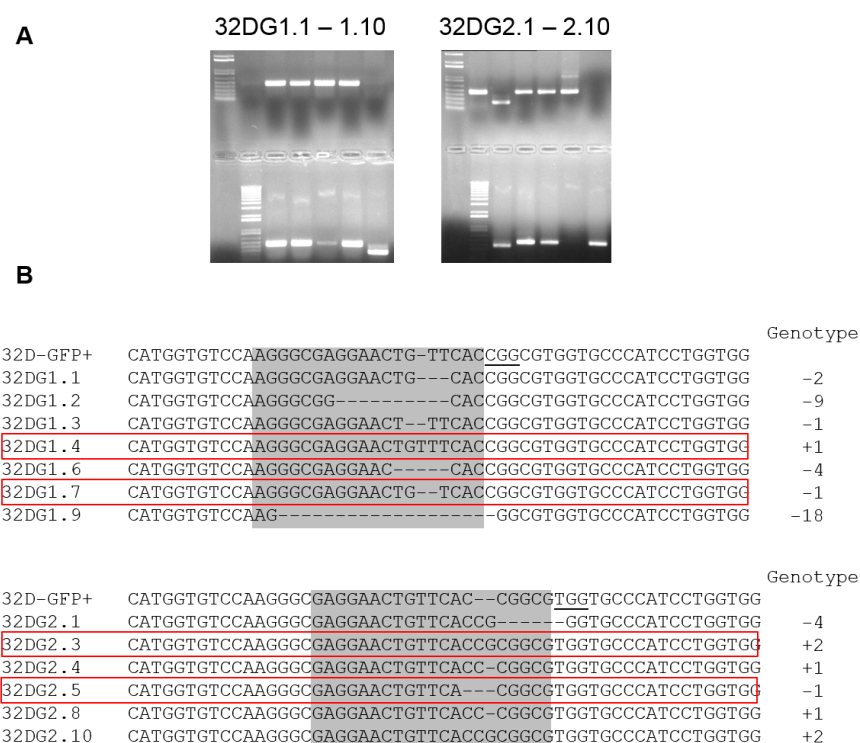


Figure 14: Indels induced by NHEJ in 32D-GFP negative cells enriched by FACS and subcloned by limiting dilution

A. EGFP PCR products (see Figure 2B) from 32D-GFP-subclones. **B.** Sequence analysis of the indels recovered by PCR.

Next, sgRNAs targeting the 32DG1.4 (+1 nt), 32DG1.7 (-1 nt), 32DG2.3 (+2 nt) and 32DG2.5 (-1 nt) mutations (Figure 14B) were nucleofected along with spCas9 into the respective clones. Consistent with previous genome editing protocols involving transfection, gene repair efficiency was low, ranging from 2.3% to 5.5% (Figure 15A).

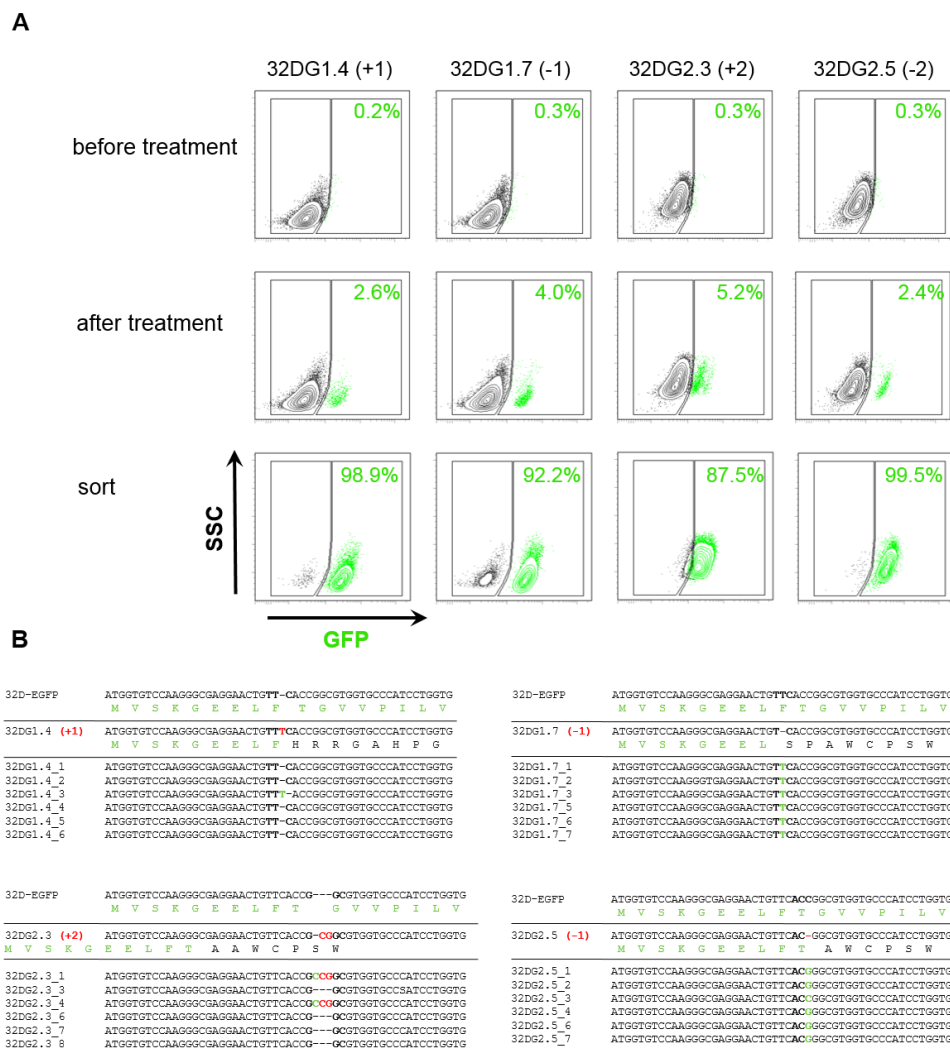


Figure 15: EGFP repair frequency in selected 32DG subclones

A. FACS and sort analysis of EGFP repair in the 32DG subclones. **B.** Indels recovered from individual subclones after sorting for EGFP expression.

To determine the type of the secondary indels leading to EGFP repair, genomic EGFP amplification products from EGFP⁺ cells of four FACS sorted 32DG clones (Figure 15A), were shot-gun cloned into the pGEM-T vector. Eight bacterial clones obtained after pGEM-T transformation of *E. coli* were isolated and sequenced from each clone. Figure 15B shows that in most cases the original EGFP reading frame was reconstituted by either frame-shifting nucleotide deletions (clones 32DG1.4 and 32DG2.5) or by frame-reconstituting nucleotide insertions (clones 32DG1.7 and 32DG2.3). However, in two instances the ORF was reconstituted by one nucleotide insertion next to the frame-shifting nucleotide (32DG1.4_3) or by one nucleotide insertion upstream of two frame-shifting nucleotides (32DG2.3_1 and 32DG2.3_4)

resulting in synonymous substitution and amino acid acquisition, respectively (Figure 15B).

3.4 NHEJ mediated reconstitution of a mutated EGFP reporter transgene in PLB-985 human promyelocytic leukemia cells

To test the NHEJ gene repair strategy in human cells, PLB985 cells expressing blue fluorescent protein (TagBFP) along with either intact (EGFP) or mutationally inactivated EGFP (mEGFP) were generated. For this, *TagBFP* was linked to *EGFP* or *mEGFP* by an *internal ribosomal entry site (IRES)*. *TagBFP-IRES-EGFP* cassettes were cloned into self-inactivating (SIN) lentiviral vector downstream of an internal *SFFV* promoter (Figure 16A). The mEGFP mutation used in the lentiviral construct was identical to the 32DG2.3 indel (Figure 14B), which creates a *SacII* restriction site by a 2-nt insertion (Figure 16A). Two lentiviral vectors referred to as SBGW and SBmGW (Figure 16A) were used to infect PLB cells at low multiplicity of infections (MOI) to obtain single copy integrations.

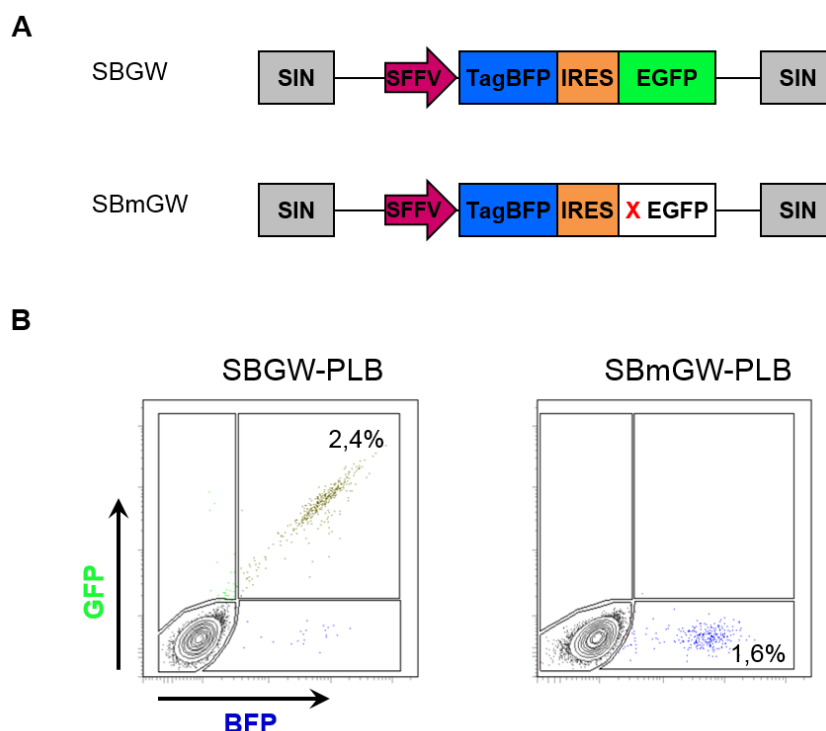
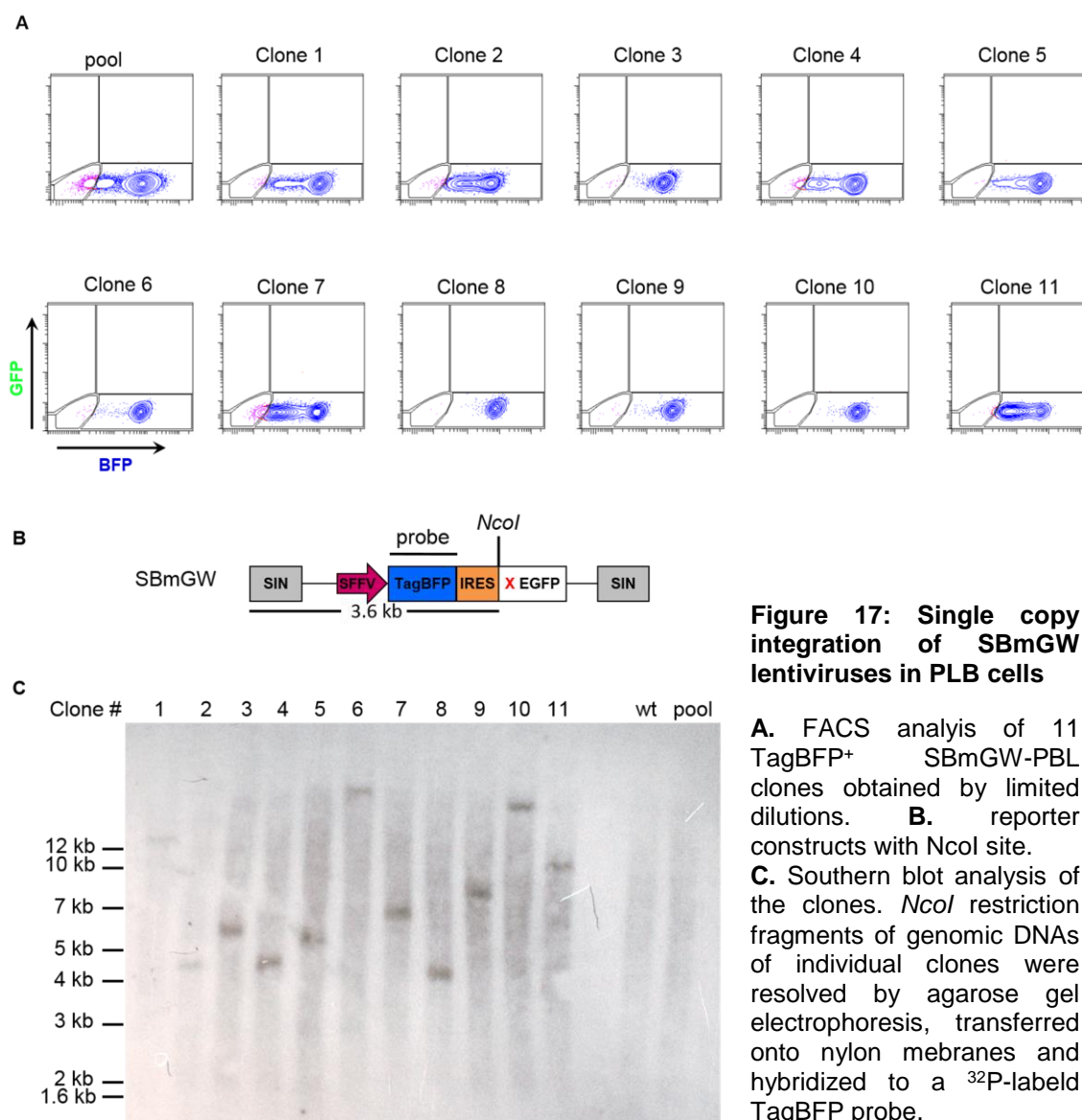


Figure 16: Generation of PLB-985 cells expressing dual color reporter transgenes

A. Lentiviral reporter constructs with cDNAs encoding blue fluorescent protein (TagBFP) and either wild type- (SBGW) or mutated (SBmGW) EGFP. **B.** FACS analysis of transduced PLB cells show that 2.4% of SBGW and 1.6% of SBmGW were transduced. Note: Only SBGW-PLB cells represent BFP/GFP double positive cells.

Two days after infection, the transduced PLBs were analyzed by FACS. As expected, the majority of SBGW transduced PLBs (SBGW-PLBs) were double positive for TagBFP and EGFP (TagBFP⁺EGFP⁺) whereas consistent with EGFP inactivation SBmGW transduced PLBs (SBmGW-PLBs) expressed only TagBFP (TagBFP⁺EGFP⁻) (Figure 16B). Transduction efficiency was 2.4% and 1.6% by SBGW and SBmGW, respectively. The SBmGW transduced cells were then sorted for TagBFP expression. To analyze SBmGW integration copy number, 11 TagBFP⁺ SBmGW-PLB clones were isolated by limiting dilution and examined by Southern blotting. As shown in Figure 17C all clones exhibited single copy integrations.



Next, sgRNA targeting the EGFP mutation (Figure 18A) were cloned into pLentiCRISPRv2 lentiviral vector¹⁴³. The construct was used to infect FACS sorted TagBFP⁺ SBmGW-PLBs by integration defective lentiviruses referred to as IDLV-LC-sgEGFP2.3 (Figure 18A). As shown in Figure 18B, infection of TagBFP⁺ SBmGW-PLBs with LC-sgEGFP2.3 reconstituted EGFP expression in up to 24% of the cells as early as 4 days after infection. Moreover, the fraction of TagBFP⁺EGFP⁺ SBmGW-PLBs slightly increased with the time in tissue culture up to 27%, indicating that EGFP repair is stable (Figure 18B). This was confirmed by Western blot showing EGFP expression in SBmGW-PLBs (Figure 18C).

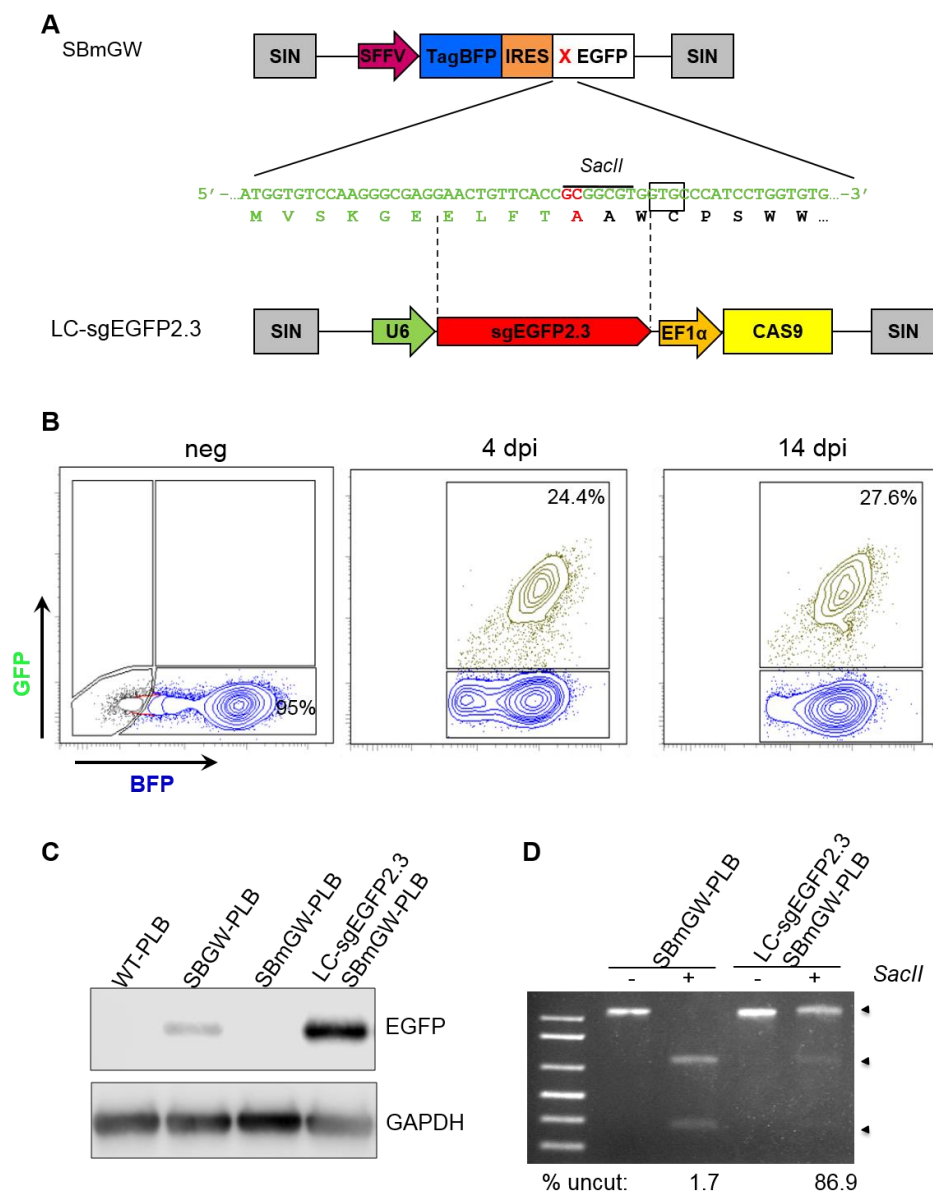


Figure 18: EGFP repair efficiency in PLB-985 cells expressing dual color reporters

A. Schematic representation of the LC-sgEGFP2.3 lentiviral vector with its target sequence. **B.** Percentage of EGFP positive cells among FACS sorted BFP⁺ SBmGW PLB cells before and after infection with integrase deficient (IDLV) LC-sgEGFP2.3 lentivirus. **C.** Western blot showing EGFP expression in WT-, unsorted SBGW-PBL control cells and in sorted TagBFP⁺ SBmGW-PLB cells before and after IDLV infection. **D.** SacII restriction fragments of genomic EGFP amplification products from SBmGW-PLB cells before and after IDLV treatment. Numbers at the bottom represent the amount of uncut DNA estimated by densitometry.

To estimate the on-target mutation rate of LC-sgEGFP2.3, genomic EGFP amplification products from the transduced SBmGW-PLBs were digested with *SacII* restriction enzyme. Following restriction fragment separation on agarose gels the uncleaved DNA was quantified by densitometry using Quantity One 1-D analysis software (Bio-Rad). As shown in Figure 18D, up to 87% of the EGFP alleles lost the *SacII* restriction site, which is consistent with a high IDLV-transduction rate. However, this on-target mutation rate is likely an underestimate as several indels recovered by shot gun cloning regenerated the *SacII* restriction site (Figure 19A). In combination with the 27% EGFP repair rate determined by FACS (Figure 18B), the indel frequency suggests that nearly every third indel restored EGFP function.

To investigate this in more detail, genomic EGFP amplification products were shot-gun cloned into pGEM-T vector as described above. Sequence analysis of bacterial clones revealed that 13 out of 28 cloned indels (46%) restored the ORF. Four of these were two nucleotide deletions (Figure 19A, C2, C23, C33 and C34) of which three restored the wild type sequence, whereas one (C33) converted Val codon into Ala (Figure 19A). The other restored ORFs included amino acid substitutions combined with acquisitions and induced by one nucleotide insertions (Gly to Ala and Ser in C16 and C23 or Gly to Ala and Gly in C18) or amino acid deletions and substitutions induced by 5 nucleotide deletions (Gly deletion in C6 and C22 or Gly + Val deletion substituted by Ala in C29) (Figure 19A). Importantly, none of the recovered sequences matched the original SBmGW sequence, suggesting an on-target mutation rate of 100%.

		Genotype	ORF	GFP
SBGW	ATGGTGTCCAAGGGCGAGGAAGTGTTCACCG---GCGTGGTGCCCATCCTGGTGGAGCTG M V S K G E E L F T G V V P I L V E L	EGFP	+	+
SBmGW (2+)	ATGGTGTCCAAGGGCGAGGAAGTGTTCACCGCG-GCGTGGTGCCCATCCTGGTGGAGCTG M V S K G E E L F T A A W C P S W W S	mEGFP	-	-
C1	ATGGTGTCCAAGGGCGAGGAAGTGTTC-----GCGTGGTGCCCATCCTGGTGGAGCTG	-6	-	-
C2	ATGGTGTCCAAGGGCGAGGAAGTGTTCACCG---GCGTGGTGCCCATCCTGGTGGAGCTG	-2	+	+
C4	ATGGTGTCCAAGGGCGAGGAAGTGTTC-----CCATCCTGGTGGAGCTG	-13	-	-
C5	ATGGTGTCCAAGGGCGAGGAAGTGTTC-----GCGTGGTGCCCATCCTGGTGGAGCTG	-6	-	-
C6	ATGGTGTCCAAGGGCGAGGAAGTGTTCAC-----CGTGGTGCCCATCCTGGTGGAGCTG	-5	+	-
C8	ATGGTGTCCAAGGGCGAGGAAGTGTTCACCGC-GCGTGGTGCCCATCCTGGTGGAGCTG	-1	-	-
C10	ATGGTGTCCAAGGGCGAGGAAGTGTTCACCG---GCGTGGTGCCCATCCTGGTGGAGCTG	-3	-	-
C11	ATGGTGTCCAAGGGCGAGGAAGTGTTCAC-----CATCCTGGTGGAGCTG	-14	+	-
C12	ATGGTGTCCAAGGGCGAGG-----TGCCCATCCTGGTGGAGCTG	-20	+	-
C13	ATGGTGTCCAAGGGCGAGGAAGTGTTCAC-----GCGTGGTGCCCATCCTGGTGGAGCTG	-48	-	-
C14	ATGGTGTCCAAGGGCGAGGAAGTGTTCACCGCG--GCGTGGTGCCCATCCTGGTGGAGCTG	-1	-	-
C15	ATGGTGTCCAAGGGCGAGGAAGTGTTCACC-CG-GCGTGGTGCCCATCCTGGTGGAGCTG	-1	-	-
C16	ATGGTGTCCAAGGGCGAGGAAGTGTTCACCGCGAGCGTGGTGCCCATCCTGGTGGAGCTG	+1	+	+
C18	ATGGTGTCCAAGGGCGAGGAAGTGTTCACCGCGGCGTGGTGCCCATCCTGGTGGAGCTG	+1	+	+
C19	ATGGTGTCCAAGGGCGAGGAAGTGTTC-----GCGTGGTGCCCATCCTGGTGGAGCTG	-269	-	-
C20	ATGGTGTCCAAGGGCGAGGAAGTGTTC-----TGCGTGGTGCCCATCCTGGTGGAGCTG	-10	-	-
C21	ATGGTGTCCAAGGGCGAGG-----TGCCCATCCTGGTGGAGCTG	-20	+	-
C22	ATGGTGTCCAAGGGCGAGGAAGTGTTCAC-----CGTGGTGCCCATCCTGGTGGAGCTG	-5	+	-
C23	ATGGTGTCCAAGGGCGAGGAAGTGTTCACCG---GCGTGGTGCCCATCCTGGTGGAGCTG	-2	+	+
C24	ATGGTGTCCAAGGGCGAGGAAGTGT-----GCCCATCCTGGTGGAGCTG	-15	-	-
C26	-----TGCGTGGTGCCCATCCTGGTGGAGCTG	-566	-	-
C27	ATGGTGTCCAAGGGCGAGGAAGTGTTCACCGCGGCGTGGTGCCCATCCTGGTGGAGCTG	+1	+	+
C28	-----GCGTGGTGCCCATCCTGGTGGAGCTG	-323	-	-
C29	ATGGTGTCCAAGGGCGAGGAAGTGTTCACCGCG-----GTGCCCATCCTGGTGGAGCTG	-5	+	-
C30	ATGGTGTCCAAGGGCGAGGAAGTGTTC-----GCGTGGTGCCCATCCTGGTGGAGCTG	-6	-	-
C33	ATGGTGTCCAAGGGCGAGGAAGTGTTCACCG---GCGTGGTGCCCATCCTGGTGGAGCTG	-2	+	+
C34	ATGGTGTCCAAGGGCGAGGAAGTGTTCACCG---GCGTGGTGCCCATCCTGGTGGAGCTG	-2	+	+
C35	ATGGTGTCCAAGGGCGAGGAAGTGTTCACCGCG---TGCGTGGTGCCCATCCTGGTGGAGCTG	-3	-	-

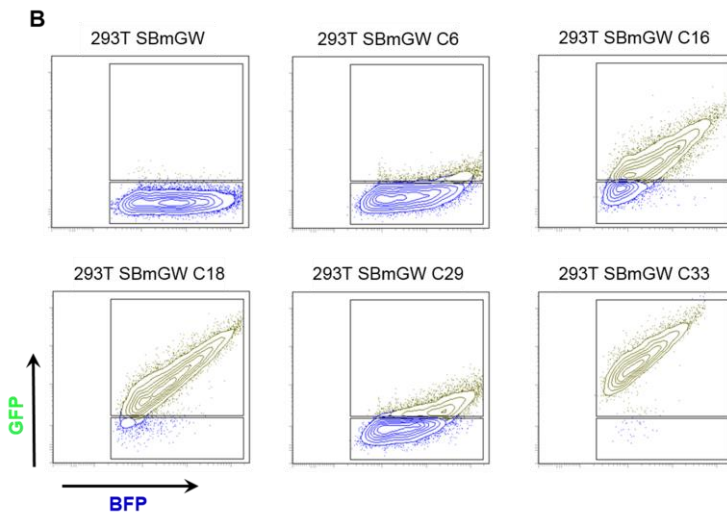


Figure 19: mEGFP alleles in IDLV infected SBmGW-PLB cells

A. Indel sequences recovered by shot-gun cloning. Reconstituted SacII restriction sites are underlined. **B.** FACS analysis of HEK293T cells expressing mEGFP cDNAs reconstituted by non-canonical ORFs.

To test whether the non-canonically reconstituted ORFs are compatible with EGFP function, the respective ORFs were replicated in SBGW vector by site-specific mutagenesis and individually transfected into HEK293T cells. FACS analysis revealed that the C16, C18 and C33 mutations were compatible with EGFP function whereas C6 and C29 were not (Figure 19B). Overall, 25% of the indels (7 out of 28) restored EGFP function (Figure 19B), which is similar to the fraction of EGFP positive cells recovered from the LC-sgEGFP2.3 transduced SBmGW-PLB cells (Figure 18B).

3.5 Reconstitution of point mutated *CYBB* gene expressed in X-CGD-PLB leukemia cells

To test whether the donor template free IDLV-RGN strategy would also effectively correct bon fide disease mutations, the EGFP in the SBGW vector was replaced with either wild type- or mutated *CYBB* cDNAs. The *CYBB* mutations were selected from the X-CGD patient specific mutation database (CYBBbase variation browser; <http://structure.bmc.lu.se/idbase/CYBBbase>). Altogether, 5 TagBFP-IRES-CYBB lentiviral vectors were generated carrying either wild type CYBB (SBwtCW) or one of the following point mutations: frameshift-R54fsCYBB (SB54CW), frameshift-R173fsCYBB (SB173CW), nonsense-E124XCYBB (SB124CW) or missense-L45RCYBB (SB45CW).

Table 6: Selected *CYBB* disease mutations and nucleotide sequences of the *CYBB*-targeting sgRNAs

Mutation	Mutation type	DNA change	Protein alteration	Accession number	sgRNA sequence
SB54CW	frameshift (duplication)	c.159dupC	p.R54fs	A53X102	CAGCACTGGCACTGGCCAGGG
SB173CW	frameshift (deletion)	c.517delC	p.L173fs	L173X188	GTACCTGGCTGTGACC-TGTGG
SB45CW	missense	c.134T>G	p.L45R	L45R	TTTACACAAGAAAACGCTCTGG
SB124CW	nonsense	c.370G>T	p.E124X	E124X	ACATCTATTTAATGTGTAATGG

As described for SBGW and SBmGW, wild-type and mutant *CYBB* carrying lentiviruses were transduced into *CYBB* null PLB985 cells (XCGD cells)¹⁴⁴ by low MOI infection to obtain single copy integrations. Two days post infection, the transduced XCGD cells were analyzed for TagBFP and *CYBB* expression by flow cytometry. As expected, the majority of SBwtCW transduced XCGD cells (XCGD-SBwtCW) expressed both TagBFP and *CYBB* whereas all XCGD cells carrying mutant copies (XCGD-SB45CW, XCGD-SB54CW, XCGD-SB124CW and XCGD-SB173CW) expressed only TagBFP (Figure 20). This was confirmed in FACS sorted TagBFP+ cells Western blot.

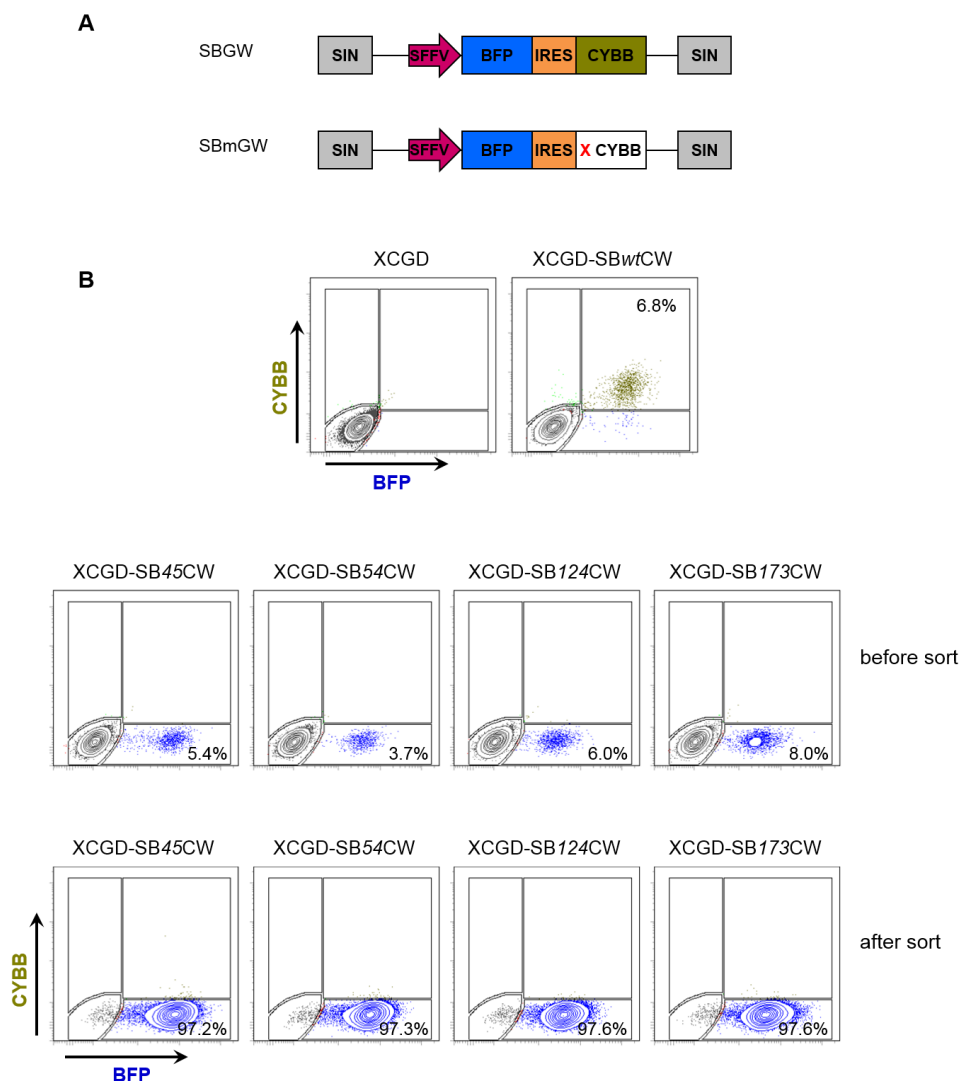


Figure 20: FACS profiles of XCGD cells expressing wild-type and mutant *CYBB* cDNAs

A. Lentiviral reporter constructs with cDNAs encoding TagBFP and either wild type- (SBwtCW) or mutated (SBmCW) *CYBB*. **B.** FACS profiles of PLB985-XCGD cells expressing wild-type or mutant *CYBB*. Note that only XCGD-SBwtCW cells are double positive for TagBFP and *CYBB*.

While SBwtCW expressed both glycosylated and non-glycosylated *CYBB*, *CYBB* expression was abolished in SB54CW, SB124CW and SB173CW XCGD cells. Interestingly, cells carrying the missense SB45CW mutation expressed only the non-glycosylated *CYBB* (Figure 21).

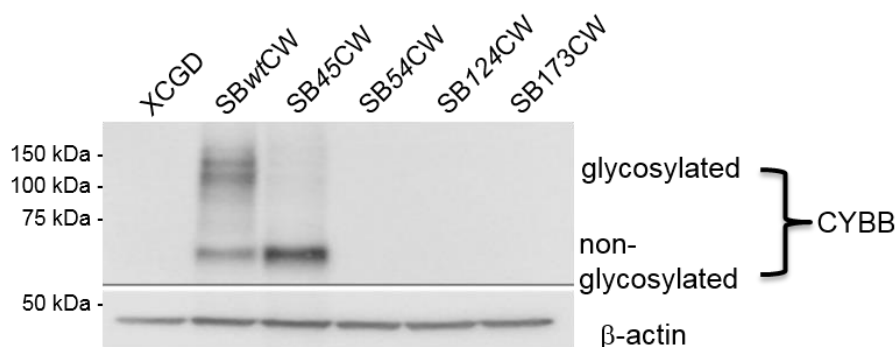


Figure 21: CYBB protein expression in XCGD-SBwtCW and SBmCW cells detected by Western blot

No CYBB expression could be detected in the PLB985-XCGD and the cells with mutated CYBB. Only the cells carrying the missense mutation in *CYBB* (SB45CW) express non-glycosylated CYBB. XCGD-SBwtCW control cells express both glycosylated and non-glycosylated CYBB.

Next, sgRNAs targeting different *CYBB* mutations (Table 6) were cloned into pLentiCRISPRv2ΔPuro and used to infect FACS-sorted, TagBFP positive XCGD cells with corresponding LC-sgCYBB IDLVs. After 14 days, up to 10% of XCGD-SB54CW- and XCGD-SB173CW cells stained positive for CYBB (TagBFP⁺CYBB⁺; Figure 22A & C). Although repair efficiency was only half of that achieved for mEGFP in SBmGW-PLB cells (Figure 18C), so was the on-target mutation rate estimated, by the Surveyor assay (Figure 22C). However, less than 2% of XCGD-SB124CW and XCGD-SB45CW cells stained positive for CYBB

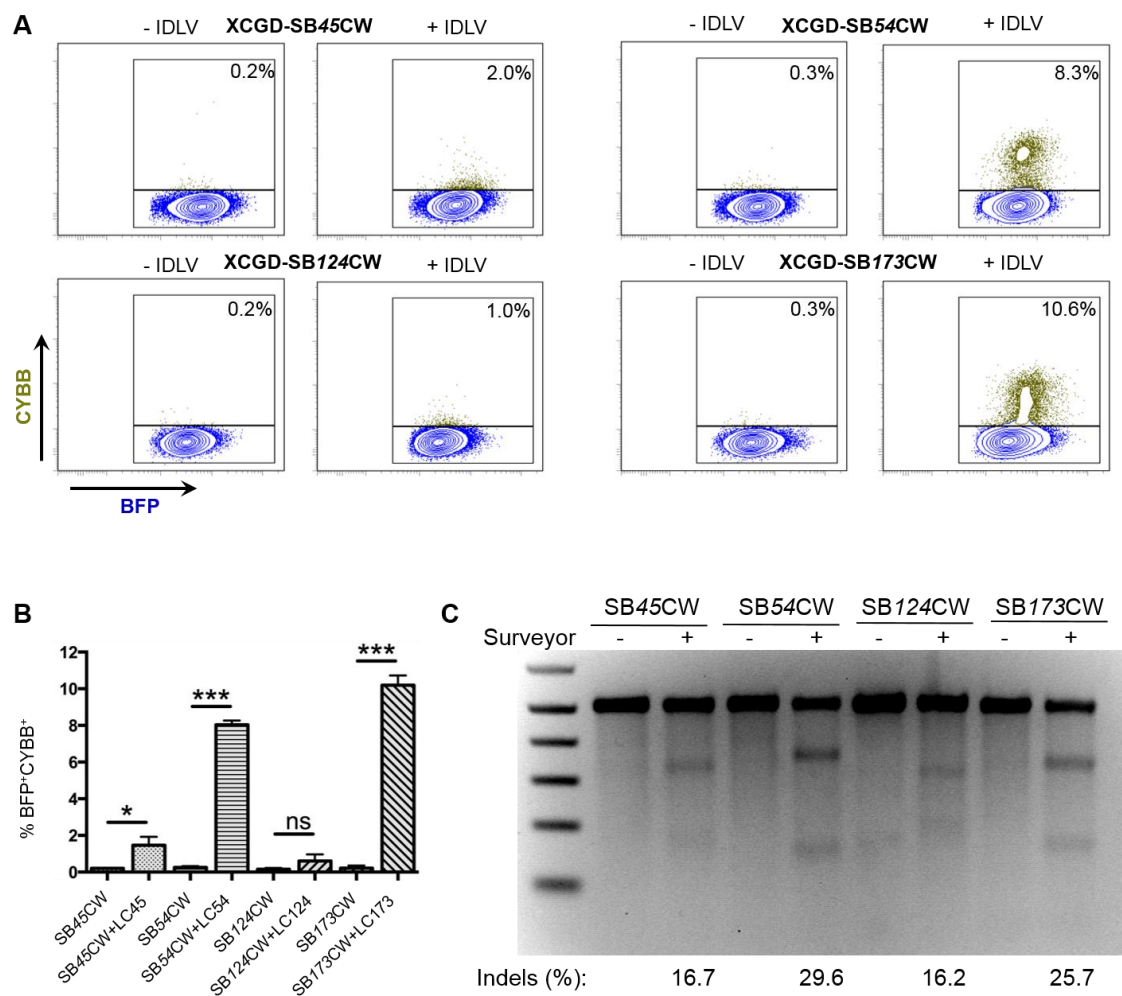


Figure 22 Repair of CYBB mutations in XCGD-PLB cells

A. FACS profiles of sorted TagBFP⁺ XCGD cells after IDLV transduction (MOI 11) and staining with the monoclonal 7D5 anti-human CYBB antibody. **B.** Frequency of CYBB positive cells amongst TagBFP⁺ XCGD cells after IDLV treatment. Results are represented as means \pm SD of 3 independent experiments. **C.** Surveyor assay of CYBB PCR products derived from IDLV treated XCGD cells. Indel frequency was calculated according to the formula published by Hsu et al. 2013.¹⁰⁵

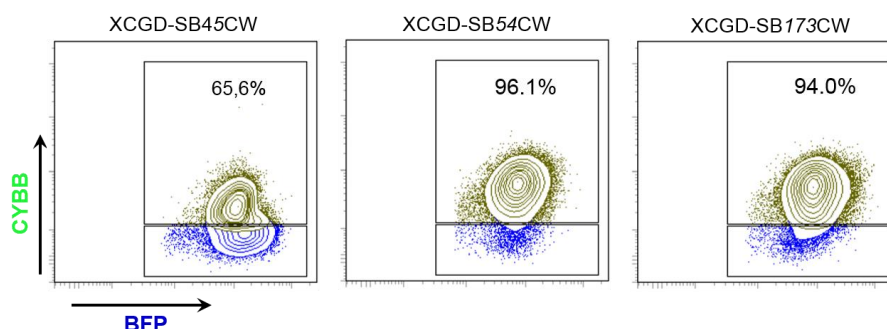


Figure 23: FACS profile of sorted CYBB⁺XCGD cells

To characterize indels compatible with CYBB expression, two different methods were used. The first one involved shot-gun cloning of CYBB amplification products from sorted, CYBB⁺ XCGD-SB45CW, XCGD-SB54CW- and XCGD-SB173CW cells (Figure 23) into pGEM-T followed by sequencing as described in section 4.4. The second involved the indel characterization in individual clones isolated by limiting dilution. As anticipated, the majority of indels isolated from the XCGD-SB45CW, XCGD-SB54CW and XCGD-SB173CW cells contained open reading frames (Table 7, Figure 24 & Figure 25).

In 22 out of 26 XCGD-SB45CW clones, the wild type CYBB sequence was reconstituted, one clone had a Lys 44 deletion and rest missense mutations of Arg 45 to His/Pro/Thr 45 (Table 7, Figure 24 & Figure 25).

Similarly, the majority of XCGD-SB54CW clones (26 out of 28) exhibited the wild type. However, two clones with single nucleotide deletions did not match the wild type sequence. (Table 7, Figure 24 & Figure 25). One clone has a R54G missense mutation and the other one has A53P modification.

In contrast, more than half of the ORFs reconstituting the XCGD-SB173CW mutation which resides in the 4th transmembrane domain of CYBB¹⁹, were non-canonical, including a 4-nt insertion, one 2-nt deletion, three 8-nt deletions and four 14-nt deletions (Table 7, Figure 24 & Figure 25). According to the SMART modular architecture research tool (<http://smart.embl-heidelberg.de/>), neither of these ORFs seriously affects the integrity of the transmembrane domain explaining why gene repair efficiency was highest in XCGD-173CW cells (Figure 22 & Table 8).

Table 7: Types of indels recovered from sorted CYBB⁺ XCGD cells

CYBB mutation	# clones	# indels	# ORFs	genotype	protein
SB45CW	41	28	26	-3 G>T(22x) G>A G>C CG>AC	K44del WT R45H R45P R45T
SB54CW	30	29	28	-1(26x) -1 -1	WT R54G A53P
SB173CW	28	23	21	+4 +1(11x) -2 -8(2x) -8 -11 -14(3x) -14	173insR WT T173_L174delinsM L173-A175del L173-G176delinsS A170-L173del V171-A175del L173-I177del

		Genotype
shot gun analysis	SB45CW	CCCCCAAGTTCTTTTACACAAGAAAAAGTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC
	C1	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>T
	C2	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>T
	C3	CCCCCAAGTTCTTTTACACAAGAAAAAG-----GGCAGCGCCCTGGCCCTGGCCAGAGCCC -4
	C4	CCCCCAAGTTCTTTTACACAAGAAAAAG-----GGCAGCGCCCTGGCCCTGGCCAGAGCCC -?
	C5	CCCCCAAGCTCTTTTACACAAGAAAAAGTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC 0
	C6	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>T & G>A
	C7	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>T
	C8	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>T
	C9	CCCCCAAGTTCTTTTACACAAGAAAAAGTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC 0
	C10	CCCCCAAGTTCTTTTACACAAGAAAAAGTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC 0
	C11	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>T
	C12	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>T & A>G & C>A
	C13	CCCCCAAGTTCTTTTACACAAGAAAAAGTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC 0
	C14	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>T
	C15	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>T & G>A
	C16	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>T
	C17	CCCCCAAGTTCTTTTACACAAGAAAAAGTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC -3
	C18	CCCCCAAGTTCTTTTACACAAGAAAAAGTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC 0
	C19	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>A
	C20	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>T
	C21	CCCCCAAGTTCTTTTACACAAGAAAAAGTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC 0
	C22	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>C
	C23	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>T
	C24	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>T
	C25	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>T
	C26	CCCCCAAGTTCTTTTACACAAGAAAAAGTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC 0
	C27	CCCCCAAGTTCTTTTACACAAGAAAAAGTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC 0
	C28	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>T
	C29	CCCCCAAGTTCTTTTACACAAGAAAAAGTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC 0
	C30	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>T
	C31	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>T
	C32	CCCCCAAGTTCTTTTACACAAGAAAAAGTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC 0
single cell analysis	C33	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>T
	C34	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>T
	C35	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>T
	C36	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>T
	C37	CCCCCAAGTTCTTTTACACAAGAAAAAGTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC 0
	C38	CCCCCAAGTTCTTTTACACAAGAAAAAGTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC 0
	C39	CCCCCAAGTTCTTTTACACAAGAAAAAGTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC 0
	C40	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC CG>AC
	C41	CCCCCAAGTTCTTTTACACAAGAAAACTTCTTGGCAGCGCCCTGGCCCTGGCCAGAGCCC G>T

		Genotype	
shot gun analysis	SB54CW	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGGCCAGGGCCCCCTGCCGCTGCCTGAAC	
	C1	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C2	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C3	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C4	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C5	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGGCCAGGGCCCCCTGCCGCTGCCTGAAC	0
	C6	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1 A>G
	C7	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C8	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C9	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C10	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C11	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C12	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C13	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C14	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C15	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C16	CCCGGAAGCTGCTGGGCTC-----CCAGGGCCCCCTGCCGCTGCCTGAAC	-15
	C17	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C18	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C19	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C20	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C21	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C22	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
single cell analysis	C23	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C24	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C25	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C26	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C27	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C28	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C29	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
	C30	CCCGGAAGCTGCTGGGCTCAGCACTGGCACTGG--CCAGGGCCCCCTGCCGCTGCCTGAAC	-1
shot gun analysis	SB173CW	CCCCGAGGGCGGCCTGTACCTGGCTGTGACC---TGTGGCCGGCATCACCGCGTGGT	0
	C1	CCCCGAGGGCGGCCTGTACCTGGCTGTGACC---TGTGGCCGGCATCACCGCGTGGT	-14
	C2	CCCCGAGGGCGGCCTGTACCTG-----GCCGGCATCACCGCGTGGT	-14
	C3	CCCCGAGGGCGGCCTGTACCTG-----GCCGGCATCACCGCGTGGT	-8
	C4	CCCCGAGGGCGGCCTGTACCTGGCTGTGA-----CCAGCATCACCGCGTGGT	-6
	C5	CCCCGAGGGCGGCCTGTACCTGGCTGTGA-----TGTGGCCGGCATCACCGCGTGGT	-8
	C6	CCCCGAGGGCGGCCTGTACCTGGCTG-----TGGCCGGCATCACCGCGTGGT	+1
	C7	CCCCGAGGGCGGCCTGTACCTGGCTGTGACC---TGTGGCCGGCATCACCGCGTGGT	+1
	C8	CCCCGAGGGCGGCCTGTACCTGGCTGTGACC---TGTGGCCGGCATCACCGCGTGGT	-8
	C9	CCCCGAGGGCGGCCTGTACCTGGCTGTGA-----CCGGCATCACCGCGTGGT	-9
	C10	CCCCGAGGGCGGCCTGTACCTG-----TGTGGCCGGCATCACCGCGTGGT	+4
	C11	CCCCGAGGGCGGCCTGTACCTGGCTGTGACCAGGCTTGTGGCCGGCATCACCGCGTGGT	0
	C12	CCCCGAGGGCGGCCTGTACCTGGCTGTGACC---TGTGGCCGGCATCACCGCGTGGT	+1
	C13	CCCCGAGGGCGGCCTGTACCTGGCTGTGACC---TGTGGCCGGCATCACCGCGTGGT	0
	C14	CCCCGAGGGCGGCCTGTACCTGGCTGTGACC---TGTGGCCGGCATCACCGCGTGGT	-2
	C15	CCCCGAGGGCGGCCTGTACCTGGCTGTGA-----TGTGGCCGGCATCACCGCGTGGT	+1
	C16	CCCCGAGGGCGGCCTGTACCTGGCTGTGACC---TGTGGCCGGCATCACCGCGTGGT	-14
	C17	CCCCGAGGGCGGCCTGTACCTGGCTGTGAC-----CACCGCGTGGT	+1
	C18	CCCCGAGGGCGGCCTGTACCTGGCTGTGACC---TGTGGCCGGCATCACCGCGTGGT	-14
	C19	CCCCGAGGGCGGCCTGTACCTG-----GCCGGCATCACCGCGTGGT	0
	C20	CCCCGAGGGCGGCCTGTACCTGGCTGTGACC---TGTGGCCGGCATCACCGCGTGGT	+1
	C21	CCCCGAGGGCGGCCTGTACCTGGCTGTGACC---TGTGGCCGGCATCACCGCGTGGT	+1
single cell analysis	C22	CCCCGAGGGCGGCCTGTACCTGGCTGTGACC---TGTGGCCGGCATCACCGCGTGGT	+1
	C23	CCCCGAGGGCGGCCTGTACCTGGCTGTGACC---TGTGGCCGGCATCACCGCGTGGT	+1
	C24	CCCCGAGGGCGGCCTGTACCTG-----TGTGGCCGGCATCACCGCGTGGT	-11
	C25	CCCCGAGGGCGGCCTGTACCTGGCTGTGACC---TGTGGCCGGCATCACCGCGTGGT	+1
	C26	CCCCGAGGGCGGCCTGTACCTGGCTGTGACC---TGTGGCCGGCATCACCGCGTGGT	+1
	C27	CCCCGAGGGCGGCCTGTACCTGGCTGTGACC---TGTGGCCGGCATCACCGCGTGGT	+1

Figure 24: Indel analyses of the shot-gun clones and single cells

Indel sequences recovered by shot-gun cloning and single cell limited dilution clones from XCGD-SB45CW, XCGD-SB54CW and XCGD-SB173CW cells. Note: Mutations out of the gRNA region were likely caused by PCR errors.

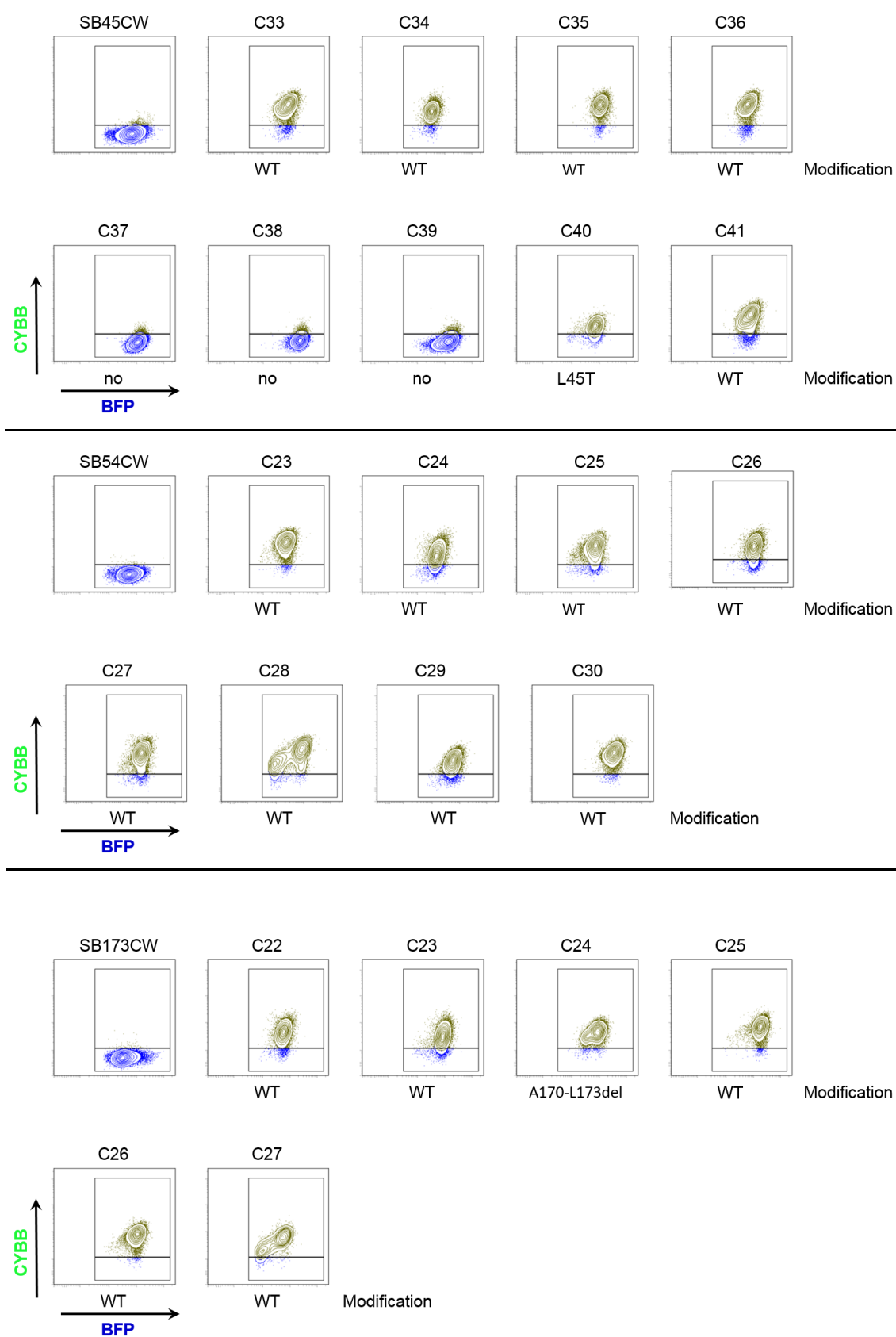
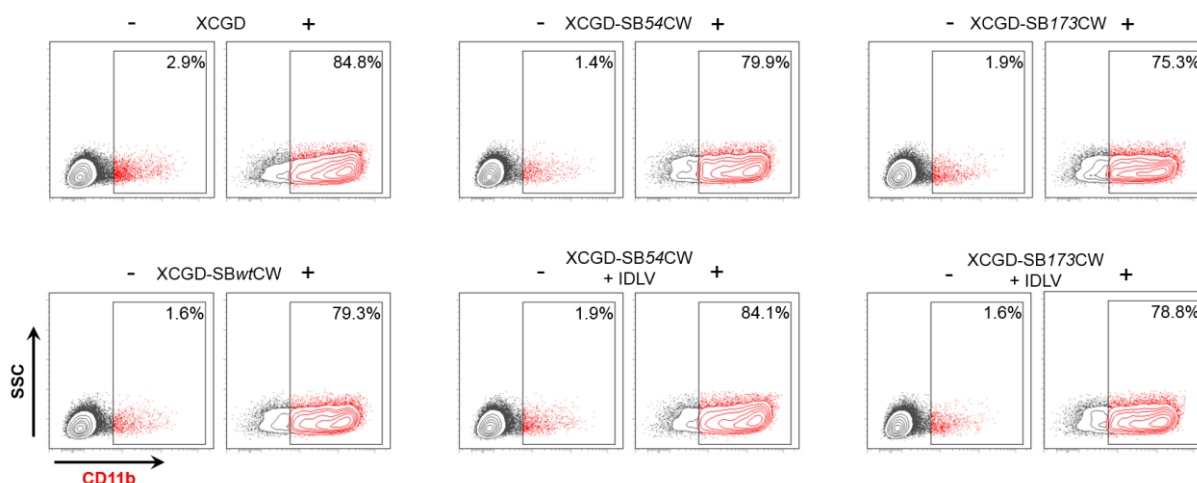


Figure 25: FACS analyses of the single cell clones from SB45CW, SB54CW and SB173CW

Table 8: Computational Transmembrane sequence analysis from XCGD-SB173CW single cell clones

CYBB mutation	genotype	protein	TM region	TM protein sequence
SB173CW	+4	173insR	169 to 191	LAVTRLLAGITGVVITLCLILII
	+1(11x)	WT	168 to 190	YLAVTLLAGITGVVITLCLILII
	-2	T173_L174delinsM	167 to 189	LYLAVMLAGITGVVITLCLILII
	-8(2x)	L173-A175del	167 to 189	GGLYLAV (A/T) GITGVVITLCLILII
	-8	L173-G176delinsS	165 to 187	GGLYLAVTSITGVVITLCLILII
	-11	A170-L173del	165 to 187	GGLYLLAGITGVVITLCLILIIIT
	-14(3x)	V171-A175del	165 to 187	GGLYLAGITGVVITLCLILIIITS
	-14	L173-I177del	165 to 187	GGLYLAVTTGVVITLCLILIIITS

To test whether CYBB⁺ XCGD-54CW- and XCGD-173CW cells exhibit NADPH oxidase activity, sorted CYBB⁺ cells (Figure 23) were induced to differentiate by DMSO and the resulting granulocytes exposed to phorbol-12-myristate-13-acetate (PMA) (Figure 26). Superoxide production was measured by using dihydrorhodamine-123 (DHR) reduction assay¹⁴². Figure 27 shows that CYBB corrected XCGD-54CW- and XCGD-173CW cells produced as much superoxide as the wild type CYBB expressing XCGD-SBwtCW cells. Moreover, as shown in Figure 27C the repaired cells expressed both non-glycosylated and glycosylated CYBB, in contrast to the untreated control cells where CYBB was undetectable.

**Figure 26: Differentiation of XCGD cells**

XCGD cells differentiation after exposure to DMSO. Cells were incubated for 7 days with DMSO (+), stained with anti-CD11b antibody and analyzed by FACS.

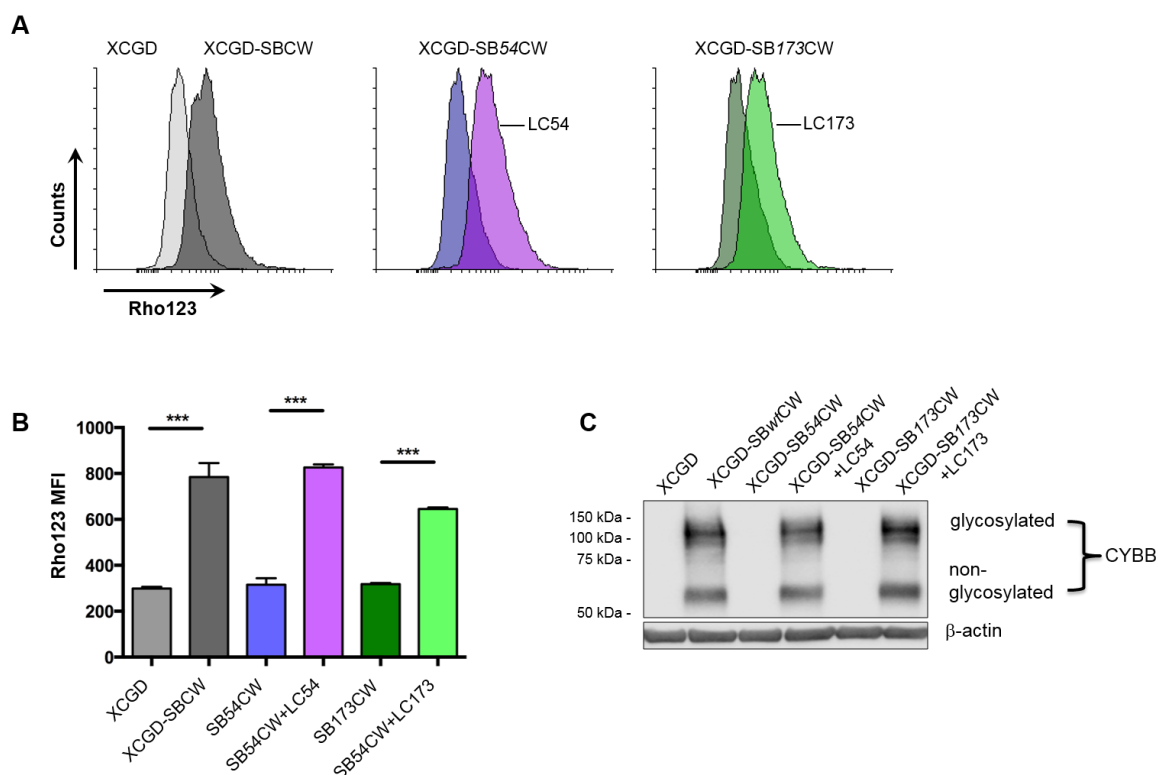


Figure 27: NADPH oxidase activity in CYBB corrected XCGD cells

A. Representative histograms depicting ROS production by differentiated, CYBB⁺ XCGD cells after stimulation with PMA. ROS levels were estimated by measuring the oxidative conversion of dihydrorhodamine123 into rhodamine123 which exhibits green fluorescence (DHR assay). **B.** Mean fluorescence intensity exhibited by differentiated XCGD cells before and after CYBB repair. XCGD and XCGD-SBwtCW cells served as negative and positive controls, respectively. **C.** Western blot showing CYBB protein expression in XCGD-SBwtCW control cells and in sorted BFP⁺ XCGD cells carrying the CYBB mutations after RGN-IDLV transduction. Results are represented as means \pm SD of 2 independent experiments. *** $p < 0.001$.

Interestingly, the clone SB173CW-C24 were get from limited dilution looks interesting, because it has 4 amino acid deletion. In silico analysis show, that the clone has 5 difference in the transmembrane domain compare to wild type CYBB (Figure 28A). To analyze, if this subclone has the same functionally activity as the pool, the subclone and the XCGD-173CW pool were analyzed by DHR-assay. Therefore subclone and the XCGD-173CW pool were differentiated to analyze NADPH activity.

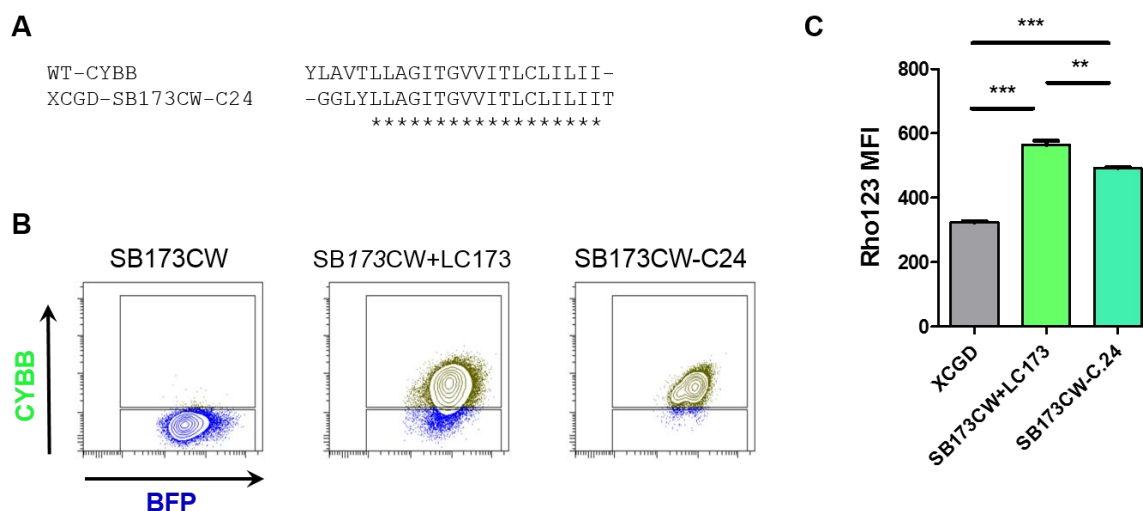


Figure 28: Functional analysis of the SB173CW-C24 clones

A. In silico analysis of the transmembrane domain in the wild type and SB173CW-C24 CYBB protein.

B. FACS data for SB173CW repair bulk and SB173CW-C24 clone. **C.** Rho123 MFI exhibited by differentiated XCGD cells as well as SB173CW repair bulk and SB173CW-C24 clone.

As shown in Figure 28, this clone were CYBB positive and the cells were Rho123-positive, suggesting a strong oxidative burst. The Rho123 MFI of subclone was only slightly lower as the pool (Figure 27C).

3.6 Estimating on target mutation efficiency at the endogenous CYBB locus in PLB cells

Finally, to determine whether the NHEJ gene repair strategy would work with similar efficiency at the endogenous *CYBB* locus, wild type PLB cells were transduced with LCBFP-sgCYBB LVs in which the sgRNAs targeting the R54fs and L173fs mutations were replaced with sgRNAs targeting the corresponding wild type sequence (sgEx3 and sgEx6; Figure 29). PLB cells transduced with LVs expressing scrambled (off target) sgRNA (LC-NTC) served as positive controls. When analyzed by flow cytometry, over 75% of the LC-sgEx3- and LC-sgEx6-LV transduced cells ceased to express CYBB (Figure 29B) which was confirmed by Western blotting (Figure 29C). To directly estimate the on target mutation rate, several indels recovered by shot-gun cloning from LC-sgEx3- and LC-sgEx6-LV transduced cells were sequenced. As shown in Figure 30, 21 out of 25 exons 3 and 19 out of 25 exons 6 exhibited a mutation, suggesting in full in agreement with the flow cytometry results an on target mutation rate of up to 75% (84% for exon 3 and 76% for exon 6). This on target mutation rate should be sufficient to correct enough cells to protect X-CGD patients from microbial infection.

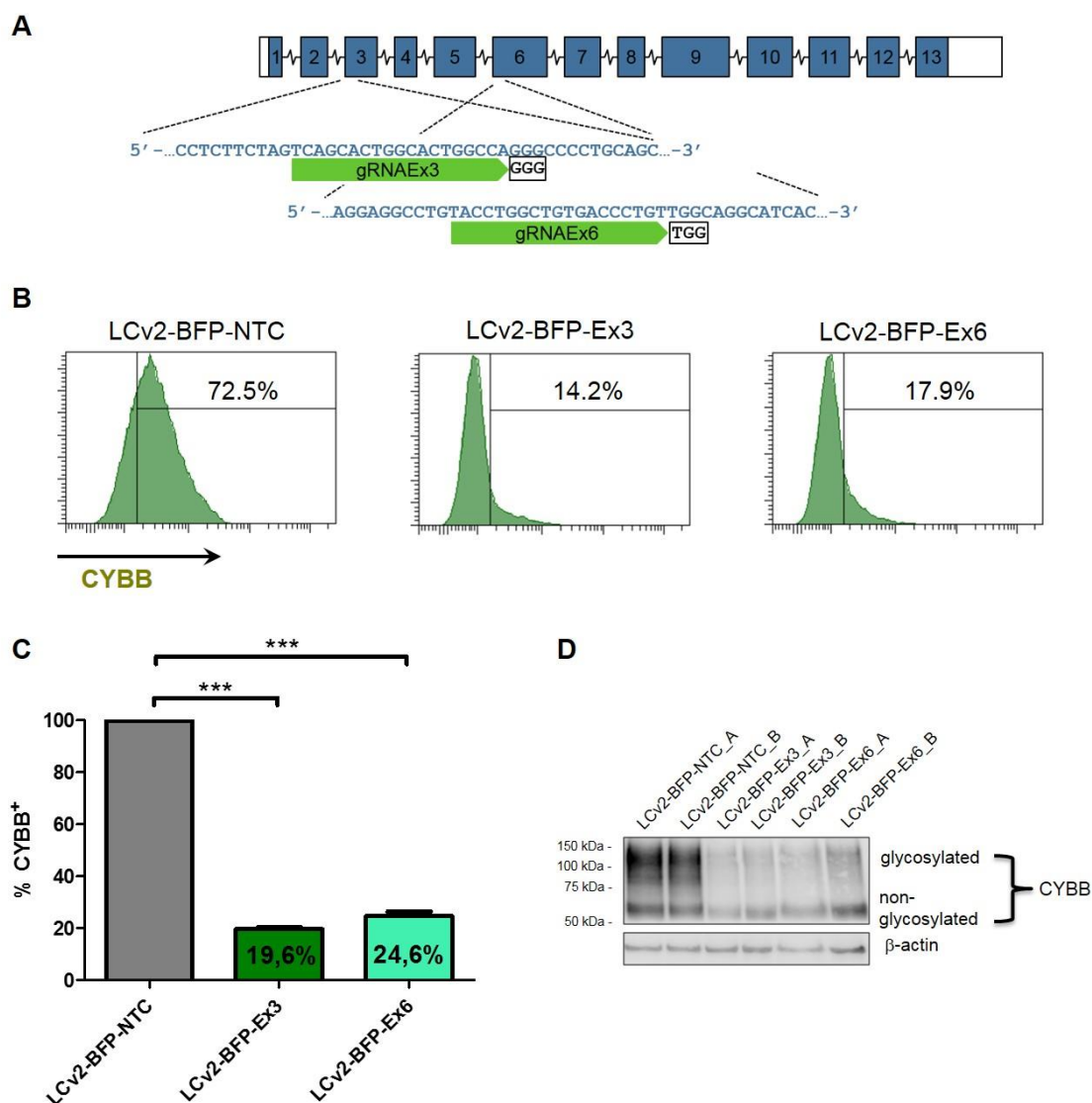


Figure 29: On target mutation efficiency at the endogenous exon 3 and exon 6 CYBB gene

A. Schematic representation of the CYBB locus with gRNA target sequences. **B.** Flow cytometry analyses of the PLB985 cells targeted with LC-sgNTC, LC-sgEx3 or LC-sgEx6. **C.** Frequency of CYBB positive cells amongst after treatment, the sgNTC treated cells was set to 100%. **D.** Western blot showing CYBB protein expression.

A

		Genotype
Ex3	CGCCTCTTCTAGTCAGCACTGGCACTGGCCAGGGCCCCCTGCAGCCTGCCTGAATTTCAAC	
C1	CGCCTCTTCTAGTCAGCACTGGCAC-----TGAGCCTGCCTGAATTTCAAC	-13
C2	CGCCTCTTCTAGTCAGCACTGGCACTG--CCAGGGCCCCCTGCAGCCTGCCTGAATTTCAAC	-1
C3	CGCCTCTTCTAGTCAGCACTGGCACTG--CCAGGGCCCCCTGCAGCCTGCCTGAATTTCAAC	-1
C4	CGCCTCTTCTAGTCAGCACTGGCACTGGCCAGGGCCCCCTGCAGCCTGCCTGAATTTCAAC	0
C5	CGCCTCTTCTAGTCAGCACTGGCACTG--CCAGGGCCCCCTGCAGCCTGCCTGAATTTCAAC	-1
C6	CGCCTCTTCTAGTCAGCACTGGCACTGGCT-----	-166
C7	CGCCTCTTCTAGTCAGCACTGG-----CCAGGGCCCCCTGCAGCCTGCCTGAATTTCAAC	-6
C8	CGCCTCTTCTAGTCAGCACTG-----CCCCTGCAGCCTGCCTGAATTTCAAC	-13
C9	CGCCTCTTCTAGTCAGCACTGGCA-----CTGCCTGAATTTCAAC	-20
C10	CGCCTCTTCTAGTCAGCACTGGCACTGGCCAGGGCCCCCTGCAGCCTGCCTGAATTTCAAC	0
C11	CGCCTCTTCTAGTCAGCACTGGCACTG--CCAGGGCCCCCTGCAGCCTGCCTGAATTTCAAC	-1
C12	CGCCTCTTCTAGTCAGCACTGGCACTG--CAGGGCCCCCTGCAGCCTGCCTGAATTTCAAC	-2
C13	CGCCTCTTCTAGTCAGCACTGGCACTGG-----TGAGCCTGCCTGAATTTCAAC	-17
C14	CGCCTCTTCTAGTCAGCACTGGCACTA--CCAGGGCCCCCTGCAGCCTGCCTGAATTTCAAC	-1
C15	CGCCTCTTCTAGTCAGCACTGGCACTG--CCAGGGCCCCCTGCAGCCTGCCTGAATTTCAAC	-1
C16	-----AC	-120
C17	CGCCTCTTCTAGTCAGCACTGGCACTG--CCAGGGCCCCCTGCAGCCTGCCTGAATTTCAAC	-1
C18	CGCCTCTTCTAGTCAGCACTGGCACTG--CCAGGGCCCCCTGCAGCCTGCCTGAATTTCAAC	-1
C19	CGCCTCTTCTAGTCAGCACTGGCAC-----TGAGCCTGCCTGAATTTCAAC	-14
C20	CGCCTCTTCTAGTCAGCACTGGCACTG--CCAGGGCCCCCTGCAGCCTGCCTGAATTTCAAC	-1
C21	CGCCTCTTCTAGTCAGCACTGGCACTGGCCAGGGCCCCCTGCAGCCTGCCTGAATTTCAAC	0
C22	CGCCTCTTCTAGTCAGCACTGGCACT--CCAGGGCCCCCTGCAGCCTGCCTGAATTTCAAC	-2
C23	CGCCTCTTCTAGTCAGCACTGGCACTGGCCAGGGCCCCCTGCAGCCTGCCTGAATTTCAAC	0
C24	CGCCTCTTCTAGTCAGCACTGGCACTG-----GCCCTGCAGCCTGCCTGAATTTCAAC	-6
C25	CGCCTCTTCTAGTCAGCACTGGCACTG--CCAGGGCCCCCTGCAGCCTGCCTGAATTTCAAC	-1

B

		Genotype
Ex6	CTGAAGGAGGCCTGTACCTGGCTGTGACCTGTGGCAGGCATCACTGGAGTTGTCATCA	
C1	CTGAAGGAG-----GCAGGCATCACTGGAGTTGTCATCA	-26
C2	CTGAAGGAGGCCTGTACCTGGCTGTGACCTGTGGCAGGCATCACTGGAGTTGTCATCA	0
C3	CTGAAGGAGGCCTGTACCTGGCTGT-----GCATCACTGGAGTTGTCATCA	-14
C4	CTGAAGGAGGCCTGTACCTGGCTGTGACCTGTGGCAGGCATCACTGGAGTTGTCATCA	0
C5	CTGAAGGAGGCCTGTACCTGG-----CAGGCATCACTGGAGTTGTCATCA	-15
C6	CTGAAGGAGGCCTGTACCTGGCTGTGACCC--GTTGGCAGGCATCACTGGAGTTGTCATCA	-1
C7	CTGAAGGAGGCCTGTACCTG-----	-47
C8	CTGAAGGAGGCCTGTACCTG-----GCATCACTGGAGTTGTCATCA	-19
C9	CTGAAGG-----TACAGGCATCACTGGAGTTGTCATCA	-27
C10	CTGAAGGAGGCCTGTACCTGGCTGTGACCTGTGGCAGGCATCACTGGAGTTGTCATCA	0
C11	CTGAAGGAGGCCTGTACCTGGCTGTGACCTGTGGCAGGCATCACTGGAGTTGTCATCA	0
C12	CTGAAGGAGGCCTGTACCTGGCTGTGACCC--GTTGGCAGGCATCACTGGAGTTGTCATCA	-1
C13	CTGAAGGAGGCCTGTACCTGT-----GTTGGCAGGCATCACTGGAGTTGTCATCA	-10
C14	-----	-198
C15	CTGAAGGAGGCCTGTACCTGG-----CAGGCATCACTGGAGTTGTCATCA	-16
C16	CTGAAGGAGGCCTGTACCTGGCTGTGACCC--ATGCAGGCATCACTGGAGTTGTCATCA	-3
C17	CTGAAGGAGGCCTGTACCTGGCTGTG-----AATCACTGGAGTTGTCATCA	-14
C18	CTGAAGGAGGCCTGTACCTGGCTGTGACCC--GTTGGCAGGCATCACTGGAGTTGTCATCA	-1
C19	CTGAAGGAGGCCTGTACCTGGCTGTGACCTGTGGCAGGCATCACTGGAGTTGTCATCA	0
C20	CTGAAGGAGGCCTGTACCTGGC-----TGTTGGCAGGCATCACTGGAGTTGTCATCA	-8
C21	CTGAAGGAGGCCTGTACCTGGCTGT-----TGTTGGCAGGCATCACTGGAGTTGTCATCA	-5
C22	CTGAAGGAGGCCTGTACCTGGCTGTGACCC--GTTGGCAGGCATCACTGGAGTTGTCATCA	-1
C23	CTGAAGGAGGCCTGTACCTGGCTGTGACCC-----GGCATCACTGGAGTTGTCATCA	-8
C24	CTGAAGGAGGCCTGTACCTGGCTGTGACCTGTGGCAGGCATCACTGGAGTTGTCATCA	0
C25	CTGAAGGAGGCCTGTACCTGGCTG-----TGTTGGCAGGCATCACTGGAGTTGTCATCA	-6

Figure 30: Indel analyses of the shot-gun clones from endogenous CYBB locus

4. Discussion

The focus of this work was to develop an RGN based strategy for *in situ* gene repair of point mutations causing monogenic blood disorders.

Over 10.000 monogenic disorders are presently known of which more than 200 are primary immune deficiencies (PIDs)¹⁰. These inherited disorders are the result of a mutation afflicting only one gene (single gene disorder) within all cells of the body. To date the only curative treatment for patients with PID is HLA-matched allogeneic hematopoietic stem cell transplantation. The probability of finding a HLA-matched donor varies among racial and ethnic groups. Patient of European descent have the best chance of finding a matched donor (75%) whereas black patients of South or Central American descent have the worst probability of finding a matched donor (16%)³⁰. For patients without a suitable HLA-matched donor, ex vivo gene therapy is the only alternative approach.

In the last three decades over 2400 clinical gene therapy trials have been initiated (April 2017: <http://www.abedia.com/wiley/phases.php>). Most trials relied on the introduction of healthy gene copies into the diseased cells using a variety of vectors including adenovirus-, adeno-associated virus (AAV)-, γ -retrovirus-, lentivirus-, and Sleeping Beauty (SB) transposon vectors. Although most strategies proved effective in terms of gene replacement, but side effects such as insertional mutagenesis leading to clonal outgrowth and leukemia's, combined with epigenetic silencing of the transgene and low engraftment rate limited their widespread use and stimulated the development of more advanced vectors with improved safety and efficacy profiles^{58,62}.

Due to these serious side effects caused by randomly integrating gene therapy vectors, novel strategies based on site-specific gene targeting and designer endonuclease mediated gene editing technologies are being pursued intensely. Accordingly, successful correction of human gene mutations by Zinc finger (ZFN) -, transcription activator like (TAL) - or by RNA guided (Cas9) endonuclease directed homology directed repair (HDR) in hematopoietic- and induced pluripotent stem cells has been reported for several inherited blood disorders including X-linked severe combined

immune deficiency (SCID)^{134,145}, sickle cell anemia^{146,147}, β -thalassemia^{148,149} and most recently, chronic granulomatous disease (CGD)^{122,150}.

In this study, designer endonucleases were used to correct point-mutated transgenes in both murine and human hematopoietic cells. To test the suitability of this approach initial studies took advantage of the *Hprt* gene expressed in an endogenous locus in 32D murine hematopoietic cell line. Despite being a “safe harbor” gene whose disruption has no functional consequences in hematopoietic cells, HPRT has the unique advantage of enabling negative and positive selection¹⁵¹. Thus, to test the efficacy of genome editing in hematopoietic cells, TAL-effector nucleases directed against the active center of *Hprt* were nucleofected into 32D cells together with a promoter-less, EGFP encoding targeting (donor) construct. After selection against *Hprt* expression, less than 5% of cells still expressed EGFP. This suggests that in the majority of the surviving cells (95%) the DSBs induced by the TAL-effector nucleases were repaired by NHEJ and only a minority (5%) by HDR (Figure 10). This was confirmed in single cell clones isolated by limiting dilution of which none underwent homologous recombination.

These results are consistent with previous studies showing that DSBs repair by NHEJ vastly exceeds DSB repair by HDR in all stages of the cell cycle^{152,153}. In contrast to NHEJ, HDR is nearly absent in G1 and exhibits its highest activity only during S phase^{152,153}. Overall, this suggests that DSBs are preferentially repaired by NHEJ rather than by HDR (95% compared to only 5%). Even if the HDR strategy would be the ideal procedure to directly replace the diseased genes at their endogenous loci, due to the very low efficiency is HDR not suitable for gene therapy for monogenic disorders. However, NHEJ rather than HDR could be a more efficient approach to gene repair.

Nevertheless, successful corrections of human disease mutations in hematopoietic- and induced pluripotent stem (IPS) cells by designer endonucleases have been reported that were exclusively based on HDR (see Table 5). In most of these applications, endonucleases and templates were delivered by nucleofection or electroporation of several plasmids, each expressing one of the functional components required for gene editing^{111,122,134,145,146,149,150}. Because these procedures are highly

toxic¹⁵⁴, only a small fraction of cells survive electroporation and even fewer retain all components required for successful genome editing. In addition to these technical considerations¹⁵⁵, DSB repair by HDR is down regulated in quiescent HSC's in favor of NHEJ¹⁵⁶. These biological and technical factors reduce the number of cells with gene repaired by HDR and therefore most protocols rely on positive selection to enrich for gene-corrected cells^{122,146,147,149}. Hence, for clinical applications, a more effective genome editing approach would seem quite useful.

Initial clinical trials employing designer endonucleases involved AIDS patients¹²⁶. In these trials ZFN were used to target and inactivate the CCR5 receptor in T-cells which is required for HIV infection. By generating DSBs within the CCR5 gene of T-cells via ZNF, the CCR5 gene was knocked out after introduction of indels by NHEJ and patients became resistant to HIV infection (clinicaltrials.gov; NCT00842634, NCT02500849)¹²⁶. While these trials demonstrated that gene therapy based on NHEJ is a suitable approach, it remained unclear how many indels reconstituted the CCR5 open reading frame (ORF). While this seems unimportant in the context of AIDS since T-cells with reconstituted CCR5 would automatically be selected against, the efficiency of gene repair by NHEJ is crucial for the success of a gene therapy of inherited blood disorders and this is exactly what the present study was designed for.

In theory, approximately one third of the indels associated with NHEJ should restore the ORF disrupted by a disease mutation. This would lead to a significant number of ORF reconstitutions of which some, depending on the position and type of the original mutation, should either completely or partially recover protein function. This is supported by the *in silico* analysis (Figure 11) of indels showing the majority (77%) of ranges between 15 nt deletions and 5 nt insertions (-15 nt to +5 nt). Assuming that one third of all indels reconstitute the reading frame of a mutated gene, approximately 25% of the targeted cells should re-express a completely or partially functional protein. Expression of a (partially) functional protein in 25% of the affected cells is sufficient for treatment of the most of the PIDs and a lot of other monogenic diseases.

Proof of concept for gene correction by NHEJ has been recently obtained in patients with Duchenne's Muscular Dystrophy (DMD)^{131,157}. DMD is an X-linked inherited disorder caused by large intragenic deletions of one or more exons disrupting the

dystrophin ORF¹⁵⁷. Dystrophin covers over 2.4 megabases contains 79 exons coding for 3,685 amino acid and is the largest human genes.

Ousterout and colleagues¹³¹ could restore the ORF of the dystrophin gene in skeletal myoblasts of patients with exons 48-50 deletion (Δ 48-51) resulting in a premature stop codon in exon 51. Using TALENs, they introduced indels into exon 51 and reconstituted the dystrophin gene in one clone. Using patients cells with the same mutation (Δ 48-50), Moggio et al.¹⁵⁷ pursued an alternative strategy in addition to this indel strategy. They multiplexed two gRNAs (RGN-RGN) to remove the targeted exon 51 in patient-derived muscle cells, thereby resetting the reading frame¹⁵⁷. This exon-removing strategy¹⁵⁷ works only for exons whose removal does not significantly interfere with protein function. Moreover, it is crucial that splicing into the adjacent downstream exon does not disrupt the ORF which has been observed to occur in rare cases.

Unlike the dystrophin studies, the present work addressed the direct correction of point mutations by NHEJ. Initial experiments involving the *Hprt* locus in 32D hematopoietic cells confirmed that NHEJ directed DSB repair is significantly more effective than HDR. Accordingly, 95% of TAL nuclease expressing cells obtained after negative selection (6-TG) harbored indels in the active center of HPRT (Figure 10). Moreover, exposure of a 6-TG resistant subclone with an indel-induced *Hprt*-inactivating frameshift mutation to a second round of TALEN-directed DSB repair yielded HAT-resistant cells derived from two independent clones - one with a synonymous substitution restoring the Ala⁵⁰ codon and one with an Arg⁵¹ to Gly⁵¹ codon conversion that only partially inactivates HPRT (Figure 12). In human pathology this mutation is associated with the Kelley-Seegmiller Syndrome¹⁵⁸. Affected patients develop urinary tract stones and mild neurological symptoms. Because this targeted mutation was in the active center of HPRT, not many modifications were accepted. For mutations located in other areas not as important to protein function, NHEJ repair will presumably accept more changes such as deletions or insertions of amino acids and thus be more efficient.

Overall, these results confirm that *in situ* gene repair by designer endonucleases does neither require donor templates nor the restoration of the original wild type sequence. However, these experiments have so far yield no information about the repair efficiency. To determine the efficiency of gene repair by NHEJ, a simple experiment

that does not require a selection and has no positive or negative effect on cell growth is sufficient. A fluorescent protein with the point mutation would be ideal for gene repair by NHEJ. 32D cells expressing EGFP from the Hprt locus after HPRT-TALEN and EGFP donor treatment (32D-EGFP⁺) were enriched and subjected to CRISPR/Cas9 mutagenesis using two different sgRNAs directed against the N-terminus of EGFP (Figure 13). Here, TALENs were abandoned after publication of CRISPR/Cas9-based gene editing. The reason to switch of strategy from TALEN to CRISPR/Cas9 system was the generally poorer efficiency of TALENs and the fact that TALENs cannot cut methylated DNA^{105,159}. CRISPR/Cas9 system is easier to design and to clone, is more cost and time-effective, usually has a high efficiency and the endonuclease can cut inside both methylated and non-methylated DNA¹⁰⁵.

Four subclones (Figure 14) with defined frameshift mutations were isolated from the mutant pools (Figure 15). T Subclones were chosen by first taking two subclones from each of the used sgRNA and second by selecting two subclones with the insertions and two with the deletions of one or two nucleotides. The clones were then nucleofected with dedicated RGNs together with Cas9 vector (two plasmid system). Gene repair efficiencies were low ranging from 2 - 5% (Figure 15) which reflected the overall toxicity of the electroporation gene delivery approach¹⁵⁴. The sequences of corrected functional EGFP shows that in the most of the cases, the EGFP ORF was restored to its original sequence by either frame-shifting the nucleotide deletions or by frame-reconstituting nucleotide insertions. However, this experiment showed that the gene repair efficiency by NHEJ without any selection can be as high as the one in the case of HDR (Figure 10) with selection. Additionally, NHEJ doesn't need any selection, meaning that it doesn't require long incubation times of HSC in clinical applications. This would probably lead to an efficient transplantation of the HSC, which undergo less differentiation *in vitro*.

The next line of research focused on adopting these processes to human cells. PLB985 human acute myeloid leukemia cell line is well-suited for this purpose, because different gene therapy approaches have demonstrated that this cell line is a powerful tool for functional analysis of the NADPH oxidase complex. Additionally, a *knock-out* CYBB model of this cell line (XCGD-PLB985 cell line) mimics a natural mutation occurring in the X-linked Chronic Granulomatous Disease(X-CGD)^{142,144,160}. In the first

experiment, PLB leukemia cells with one copy of a mutant EGFP reporter inserted into their genome were used to test the NHEJ approach in human cells. Using template free RGNs delivered by integrase-defective lentiviruses (IDLVs), on-target mutation rates approaching 100% were obtained (Figure 18 & Figure 19). Combined with the achieved gene repair efficiency of 27% (Figure 18, section 3.4), the result suggests that about one in four NHEJ events reconstituted EGFP function. This repair efficiency exceeded the EGFP repair efficiency in 32D cells by almost 5- to 10-fold. The majority of these effects can be attributed to the IDLV–RGN delivery approach. This system, unlike the two plasmid system employed by nucleofection, delivers all components required for gene editing within one vector by cell infection, - an approach generally less toxic than electroporation. Unlike IDLVs which engage cell surface receptors for cell entry, electroporation requires cell permeabilization leading to cell content leaks adversely affecting adjacent cells. Moreover, electroporation generates random DSBs that can cause mutations^{161,162}. Additionally, transfected donor DNA after inserting can integrate randomly into the genome after insertion. IDLVs express the RGNs episomally from one or two LTR circles which are diluted out by cell replication ensuring transient expression and preventing insertional mutagenesis^{163–167} – a major safety issue in clinical trials employing integrating lentiviruses. Although between 0.35 and 2.3% of IDLVs can still integrate into the genome, depending on the IN mutation and vector dose^{165,168} the probability of long term adverse effects is rather low because constitutive Cas9 expression activates the immune system leading to elimination of Cas9 expressing cells¹⁶⁹. Remaining safety concerns with IDLVs could be addressed by inserting suicide genes such as the herpes simplex virus thymidine kinase (HSV-tk) gene enabling the elimination of HSV-tk expressing cells in vivo by administering the pro-drug as ganciclovir^{170,171}.

Analysis of the reconstituted open reading frames (ORFs) showed that all modification with single amino acid deletions were compatible with EGFP fluorescence (Figure 19). This is consistent with recent reporting showing that loss of amino acids from the N-terminal α -helix of GFP negatively interfere with EGFP fluorescence¹⁷². Interestingly, analysis of the reconstituted open reading frames (ORFs) showed that all but single amino acid deletions were compatible with EGFP fluorescence. More specifically, the short hydrophobic N-terminal α -helix of EGFP stabilizes its barrel structure which is

required for fluorescence and therefore only small hydrophobic amino acids, such as alanine are tolerated at this site¹⁷²

Next, the template free IDLV-RGN strategy was tested on a real monogenic diseases. For this purpose, X-linked chronic granulomatous disease (X-CGD) were selected. As already described, there exists a cell line *CYBB* knock-out (KO) model for X-CGD (XCGD-PLB985) and after *knock-in* (KI) of the *CYBB* gene, XCGD-PLB985 cells are a powerful tool for functional analysis of the NADPH oxidase complex. Importantly, neither the KO nor the KI has any effect on cell growth¹⁵. If the cells with KO or KI had a proliferation advantage, experiments assessing on-target gene editing efficiency would yield biased data.

To test the RGN strategy on *bona fide* disease genes, several patient specific *CYBB* mutations causing X-CGD were selected from the Immunodeficiency mutation databases (*CYBB*base variation browser; <http://structure.bmc.lu.se/idbase/CYBBbase>), including two frameshift-, one missense- and one nonsense mutation. *In vitro* models for X-CGD were obtained by introducing the mutations as *CYBB* transgenes into *CYBB*-KO PLB cells (XCGD-PLB985)¹⁴⁴. As promyelocytic leukemia cells expressing the NADPH oxidase complex, PLB cells were perfectly suited for testing the RGN approach¹⁵. Treatment of the respective mutations with dedicated RGNs showed that up to 10% of the transduced XCGD cells harbouring *CYBB* frameshift mutations recovered *CYBB* expression and produced superoxide anions after differentiation into granulocytes. Considering the estimated on-target mutation rates of 30% and 26% in XCGD-SB54CW- and XCGD-SB173C cells, respectively (Figure 22) the actual frequency of gene repairing NHEJ events is estimated to be 27% for the SB54CW - and 38% for the SB173CW mutation (Table 9).

In comparison, both nonsense and missense *CYBB* mutations showed reduced on-target mutation rates (Figure 22) which may be due in part to the low CG content of the respective sgRNAs (Table 9)¹⁷³. Moreover, as Cas9 has been shown to tolerate single nucleotide mismatches¹⁷⁴, selection against single nucleotide substitutions by RGN retargeting of already corrected templates may reduce repair efficiency even further.

Table 9: Theoretical frequency of gene repair

Mutation	Mutation type	On target [%]	Repair [%]	Theoretical repair efficiency by 100% on target [%]	GC content
SB54CW	frameshift (duplication)	29.6	8.0	27.0	68
SB173CW	frameshift (deletion)	25.7	10.3	40.1	58
SB45CW	missense	16.7	1.8	10.8	32
SB124CW	nonsense	16.2	0.8	4.9	21

$$\text{Theoretical repair efficiency} = \text{Repair} * \frac{100}{\text{On Target}}$$

However, considering that residual NADPH oxidase activity at 5 - 10% of normal levels is already sufficient to protect X-CGD patients from microbial infections^{20,175} and that even residual activity of 1 - 2% of normal level provides a significant survival benefit^{175,176}, even the low repair rates of missense and nonsense mutations would suffice if RGN transduction rates approached 100% (Table 9). In addition, the reported on target mutation rates are likely an underestimate because the sensitivity of the employed classic surveyor assay is limited¹⁷⁷. In some cases the real mutation rate could be 20% higher than that measured by surveyor assay.

Notwithstanding the above considerations, repair rates are also contingent on the respective targeted mutation sites. Whereas the SB45CW mutation is located within a splice site of exon 2¹⁷⁸ and required reconstitution of the wild type ORF. The SB54CW mutation is located in the redox active center of CYBB^{81,179–181} almost exclusively required wild type ORF reconstitution, too. In contrast, the SB 173CW mutation located in the fourth transmembrane domain also tolerated a variety of non-canonical ORFs (Table 7 & Figure 24 & Figure 25). According to the SMART modular architecture research tool (<http://smart.embl-heidelberg.de/>), neither of these ORFs seriously affects the integrity of the transmembrane domain, thus explaining why the gene repair efficiency was highest in XCGD-173CW cells.

To determine whether the NHEJ gene repair strategy would be similarly efficient at the endogenous locus, wild type PLB cells were treated with RGNs targeting the exons harbouring the R54fs and L173fs mutations. In both cases, the KO analysis of CYBB by flow cytometry and western blot (Figure 29) shows KO efficiency above 75% and the on target mutation rate determined by sequencing of indels recovered by shot gun cloning was also above 75%. Assuming gene repair efficiencies of about a quarter (Figure 18), this data imply an *in situ* gene repair efficiency of at least 20%, which is sufficient to protect X-CGD patients from microbial infection and thereby prevent the major cause of death in X-CGD.

The next step to continue this work would have been final validation of the outlined strategy in primary X-CGD patient cells, but it was not possible to obtain suitable primary material. X-CGD is a rare disorder and most patients are severely ill, rendering access to primary material by bone marrow aspiration or apheresis following HSPC mobilization by G-CSF very challenging. Bone marrow aspiration is painful, complex and carries the risk of both patient infections and contamination of material. Apheresis is easier, but HSPC must be mobilized with recombinant granulocyte colony-stimulating factor (G-CSF), which can have side effects such as severe bone pains, headache, arthralgia, malaise, fatigue, insomnia nausea¹⁸². A suitable and consenting donor could not be found in the available time, and therefore an IDLV-RGN repair in primary human HSPCs could not be demonstrated here.

5. Conclusions and Outlook

The present work has provided proof of concept for site directed gene repair by NHEJ in an X-CGD model.

The data shown provide evidence for high efficiency gene repair by donor template free CRISPR/Cas9 technology using transient, IDLV based RGN transduction into hematopoietic cells. Unlike the widely adopted gene repair by homologous recombination (HDR), gene repair by NHEJ requires no donor templates, and is significantly more efficient because unlike HDR, operates in all phases of the cell cycle. Furthermore, NHEJ is the dominant DSB-repair pathway in quiescent (G0) HSCs^{153,155,156}, which are the desired target cells for gene therapy of most monogenic blood disorders. Procedurally, the single vector IDLV-RGN delivery made possible by a template-free CRISPR/Cas9 approach is one of its most attractive technical features, because it is transient, more efficient and less toxic than plasmid-based technologies. Thus, gene therapy based on the template free IDLV-RGNs provides a sensible alternative to existing gene therapy protocols. The present study has shown that frameshift mutations can be effectively repaired in hematopoietic cells by donor template free CRISPR/Cas9 gene editing

According to the CYBBbase database¹⁸³, 24% of all X-CGD patients harbor *CYBB* frameshift mutations. This frequency is similar to the frequency of frameshift mutations found in the *IL2R γ* ^{184,185}, *WASP*^{186,187}, *ADA*¹⁸⁸, and *HBB*^{189,190} genes of patients with X-linked immunodeficiency disease (X-SCID), Wiskott-Aldrich Syndrome, adenosine deaminase immunodeficiency disease (ADA-SCID) and β -thalassemia, respectively. Thus, at least one in four to five of these patients could benefit from gene therapy with donor template free, RNA-guided Cas9 endonucleases. For most of these genes, recover of less than 10% of wild type protein activity is already sufficient for obtaining a therapeutic effect. For immunodeficiency diseases in which positive selection occurs (e.g. X-SCID, ADA-SCID, WAS), even lower gene repair rates are sufficient because repaired cells will overgrow diseased cells over time after transplantation.

However, before reaching clinical maturity the strategy needs to be validated in vivo. For this, transgenic mice carrying humanized and patient-specific X-CGD or X-SCID mutations produced by single embryo CRISPR/Cas9 gene editing¹⁹¹ will be used for syngeneic bone marrow transplantation experiments in which ex vivo repaired hematopoietic stem cells are injected into lethally irradiated recipients followed by long term disease and gene expression monitoring. Special attention will be paid to potential adverse effects caused by off-target mutations detected by dedicated reporter assays and whole genome sequencing which will also reveal possible IDLV genome integrations.

In a second preclinical approach which is largely dependent on patient availability (see Discussion), hematopoietic stem and progenitor cells (HSPCs) recovered from the peripheral blood after G-CSF mobilization will be transplanted after ex vivo gene repair into immunodeficient NSG mice¹⁵⁰. Successful gene repair rates and off-target effects will then be monitored by standard methodology.

Ideally, RGN delivery for gene therapy should be transient and virus free to avoid insertional mutagenesis and immunological side effects. Therefore, special consideration should be given to alternative RGN delivery strategies, in the future experiments. Most promising in this regard are advanced electroporation strategies delivering the RGNs as preassembled gRNA/Cas9 protein complexes^{192,193}.

Finally, the donor template free CRISPR/Cas9 approach should be easily adaptable to non-SCID diseases such as Haemophilia or Duchenne Muscular Dystrophy.

1. References

1. Maximow, A. Der Lymphozyt als gemeinsame Stammzelle der verschiedenen Blutelemente in der embryonalen Entwicklung und im postfetalen Leben der Säugetiere. *Folia Haematol.* 125–134 (1909).
2. Rieger, M. A. and & SCHROEDER, T. Hämatopoetische Stammzellen. *BIOspektrum* (2007).
3. Rieger, M. A. & Schroeder, T. Exploring hematopoiesis at single cell resolution. *Cells Tissues Organs* **188**, 139–149 (2008).
4. Smith, C. Hematopoietic stem cells and hematopoiesis. *Cancer Control* **10**, 9–16 (2003).
5. Rieger, M. A. & Schroeder, T. Hematopoiesis. *Cold Spring Harb Perspect Biol* **4**, (2012).
6. Gunsilius, E., Gastl, G. & Petzer, A. L. Hematopoietic stem cells. *Biomed. Pharmacother.* **55**, 186–194 (2001).
7. WHO. <http://www.who.int/genomics/public/geneticdiseases/en/index2.html> at
<<http://www.who.int/genomics/public/geneticdiseases/en/index2.html>>
8. Farhan, S. M. K. & Hegele, R. a. Genetics 101 for Cardiologists: Rare Genetic Variants and Monogenic Cardiovascular Disease. *Can. J. Cardiol.* **29**, 18–22 (2013).
9. Kaufmann, K. B., Büning, H., Galy, A., Schambach, A. & Grez, M. Gene therapy on the move. *EMBO Mol. Med.* **5**, 1642–1661 (2013).
10. Nonas, S. Pulmonary Manifestations of Primary Immunodeficiency Disorders. *Immunol. Allergy Clin. North Am.* **35**, 753–766 (2015).
11. Cavazzana-calvo, M. & Fischer, A. Gene therapy for severe combined immunodeficiency : are we there yet ? *Sci. Med.* **117**, 1456–1465 (2007).
12. Al-Herz, W. *et al.* Primary immunodeficiency diseases: An update on the classification from the International Union of immunological societies expert committee for primary immunodeficiency. *Front. Immunol.* **5**, 1–33 (2014).
13. Lim, M. S. & Elenitoba-johnson, K. S. J. The Molecular Pathology of Primary Immunodeficiencies. **6**, 59–83 (2004).
14. M.L., L. *et al.* Applying public health strategies to primary immunodeficiency diseases: a potential approach to genetic disorders. *MMWR. Recomm. Rep.* **53**, 1–29 (2004).
15. Aiuti, A., Bacchetta, R., Seger, R., Villa, A. & Cavazzana-Calvo, M. Gene therapy for primary immunodeficiencies: Part 2. *Curr. Opin. Immunol.* **24**, 585–591 (2012).
16. Aiuti, A. *et al.* Lentiviral Hematopoietic Stem Cell Gene Therapy in Patients with Wiskott-Aldrich Syndrome. *Science (80-.).* **341**, 1233151 (2013).
17. Jin, Y. *et al.* Mutations of the Wiskott-Aldrich Syndrome Protein (WASP): hotspots, effect on transcription, and translation and phenotype/genotype correlation. *Blood* **104**, 4010–9 (2004).
18. Rudd, C. E. Disabled receptor signaling and new primary immunodeficiency disorders. *N. Engl. J. Med.* **354**, 1874–1877 (2006).
19. Rae, J. *et al.* X-Linked chronic granulomatous disease: mutations in the CYBB gene encoding the gp91-phox component of respiratory-burst oxidase. *Am J Hum Genet* **62**, 1320–1331 (1998).
20. Becker, S. *et al.* Correction of respiratory burst activity in X-linked chronic granulomatous cells to therapeutically relevant levels after gene transfer into bone marrow CD34+ cells. *Hum Gene Ther* **9**, 1561–1570 (1998).
21. Pessach, I. M. & Notarangelo, L. D. Gene therapy for primary immunodeficiencies: Looking ahead, toward gene correction. *J. Allergy Clin. Immunol.* **127**, 1344–1350 (2011).
22. Antonarakis, S. E., Krawczak, M. & Cooper, D. N. Disease-causing mutations in the human genome. *Eur J Pediatr* **159 Suppl**, S173-8 (2000).

23. Mercadier, J. How do mutations cause disease ? 34–37 (2008).
24. Frischmeyer, P. a & Dietz, H. C. Nonsense-mediated mRNA decay in health and disease. *Hum. Mol. Genet.* **8**, 1893–1900 (1999).
25. Kuzmiak, H. A. & Maquat, L. E. Applying nonsense-mediated mRNA decay research to the clinic: progress and challenges. *Trends Mol. Med.* **12**, 306–316 (2006).
26. Henig, I. & Zuckerman, T. Hematopoietic stem cell transplantation-50 years of evolution and future perspectives. *Rambam Maimonides Med. J.* **5**, e0028 (2014).
27. Cavazzana-Calvo, M., Fischer, A., Hacein-Bey-Abina, S. & Aiuti, A. Gene therapy for primary immunodeficiencies: Part 1. *Curr. Opin. Immunol.* **24**, 580–584 (2012).
28. Champlin, R. E. *et al.* Blood stem cells compared with bone marrow as a source of hematopoietic cells for allogeneic transplantation. IBMTR Histocompatibility and Stem Cell Sources Working Committee and the European Group for Blood and Marrow Transplantation (EBMT). *Blood* **95**, 3702–3709 (2000).
29. Gennery, A. R. *et al.* Transplantation of hematopoietic stem cells and long-term survival for primary immunodeficiencies in Europe: Entering a new century, do we do better? *J. Allergy Clin. Immunol.* **126**, (2010).
30. Gragert, L. *et al.* HLA Match Likelihoods for Hematopoietic Stem-Cell Grafts in the U.S. Registry. *N. Engl. J. Med.* **371**, 339–348 (2014).
31. Nayerossadat, N., Ali, P. & Maedeh, T. Viral and nonviral delivery systems for gene delivery. *Adv. Biomed. Res.* **1**, 27 (2012).
32. Tros de Ilarduya, C., Sun, Y. & Düzgünes, N. Gene delivery by lipoplexes and polyplexes. *Eur. J. Pharm. Sci.* **40**, 159–170 (2010).
33. Manjila, S. B. *et al.* Novel gene delivery systems. *Int. J. Pharm. Investig.* **3**, 1–7 (2013).
34. Xiang, S. *et al.* Uptake mechanisms of non-viral gene delivery. *J. Control. Release* **158**, 371–378 (2012).
35. Skipper, K., Andersen, P., Sharma, N. & Mikkelsen, J. DNA transposon-based gene vehicles - scenes from an evolutionary drive. *J. Biomed. Sci.* **20**, 92 (2013).
36. Bouuaert, C. C. & Chalmers, R. M. Gene therapy vectors: The prospects and potentials of the cut-and-paste transposons. *Genetica* **138**, 473–484 (2010).
37. Ivics, Z., Izsvák, Z., Medrano, G., Chapman, K. M. & Hamra, F. K. Sleeping Beauty transposon mutagenesis in rat spermatogonial stem cells. *Nat. Protoc.* **6**, 1521–1535 (2011).
38. Field, A.-C. *et al.* Comparison of lentiviral and sleeping beauty mediated $\alpha\beta$ T cell receptor gene transfer. *PLoS One* **8**, e68201 (2013).
39. Gogol-Doring, A. *et al.* Genome-wide Profiling Reveals Remarkable Parallels Between Insertion Site Selection Properties of the MLV Retrovirus and the piggyBac Transposon in Primary Human CD4(+) T Cells. *Mol. Ther.* **24**, 592–606 (2016).
40. Ivics, Z., Hackett, P. B., Plasterk, R. H. & Izsvák, Z. Molecular Reconstruction of Sleeping Beauty, a Tc1-like Transposon from Fish, and Its Transposition in Human Cells. *Cell* **91**, 501–510 (1997).
41. Izsvák, Z., Chuah, M. K. L., VandenDriessche, T. & Ivics, Z. Efficient stable gene transfer into human cells by the Sleeping Beauty transposon vectors. *Methods* **49**, 287–297 (2009).
42. Johnen, S. *et al.* Antiangiogenic and Neurogenic Activities of Sleeping Beauty-Mediated PEDF-Transfected RPE Cells In Vitro and In Vivo. *Biomed Res. Int.* **2015**, 863845 (2015).
43. Bouard, D., Alazard-Dany, D. & Cosset, F.-L. Viral vectors: from virology to transgene expression. *Br. J. Pharmacol.* **157**, 153–65 (2009).
44. Mali, S. Delivery systems for gene therapy. *Indian J. Hum. Genet.* **19**, 3–8 (2013).
45. Robbins, P. D. & Ghivizzani, S. C. Viral Vectors for Gene Therapy. *Pharmacol. Ther.* **80**, 35–47 (1998).
46. Gardlík, R. *et al.* Vectors and delivery systems in gene therapy. *Med. Sci. Monit.* **11**, RA110-A121 (2005).

47. Kootstra, N. A. & Verma, I. M. Gene therapy with viral vectors. *Annu. Rev. Pharmacol. Toxicol.* **43**, 413–439 (2003).
48. Deyle, D. R. & Russell, D. W. Adeno-associated virus vector integration. *Curr. Opin. Mol. Ther.* **11**, 442–7 (2009).
49. Nault, J.-C. *et al.* Recurrent AAV2-related insertional mutagenesis in human hepatocellular carcinomas. *Nat. Genet.* **47**, 1187–1193 (2015).
50. Russell, D. W. & Grompe, M. Adeno-associated virus finds its disease. *Nat. Genet.* **47**, 1104–1105 (2015).
51. Lech, P. & Somia, N. Retrovirus vectors. *Contrib. Nephrol.* **159**, 30–46 (2008).
52. Baum, C., Schambach, A., Böhne, J. & Galla, M. Retrovirus Vectors: Toward the Plentivirus? *Mol. Ther.* **13**, 1050–1063 (2006).
53. El Ashkar, S. *et al.* Engineering Next-Generation BET-Independent MLV Vectors for Safer Gene Therapy. *Mol. Ther. - Nucleic Acids* **7**, 231–245 (2017).
54. Stein, S. *et al.* From Bench to Bedside: Preclinical Evaluation of a Self-Inactivating Gammaretroviral Vector for the Gene Therapy of X-linked Chronic Granulomatous Disease. *Hum. Gene Ther. Clin. Dev.* **24**, 86–98 (2013).
55. De Ravin, S. S. *et al.* Lentiviral hematopoietic stem cell gene therapy for X-linked severe combined immunodeficiency. *Sci. Transl. Med.* **8**, 335ra57 (2016).
56. Farinelli, G. *et al.* Lentiviral vector gene therapy protects XCGD mice from acute Staphylococcus aureus pneumonia and inflammatory response. *Mol. Ther.* (2016). doi:10.1038/mt.2016.150
57. Spanevello, F. *et al.* Development of Lentiviral Vectors Simultaneously Expressing Multiple siRNAs Against CCR5, vif and tat/rev Genes for an HIV-1 Gene Therapy Approach. *Mol. Ther. Acids* **5**, e312 (2016).
58. Naldini, L. Ex vivo gene transfer and correction for cell-based therapies. *Nat Rev Genet* **12**, 301–315 (2011).
59. Mitchell, R. S. *et al.* Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol.* **2**, (2004).
60. Cattoglio, C. *et al.* High-definition mapping of retroviral integration sites identifies active regulatory elements in human multipotent hematopoietic progenitors High-definition mapping of retroviral integration sites identifies active regulatory elements in human multipotent. *Gene Ther.* **116**, 5507–5517 (2011).
61. Qasim, W., Gaspar, H. B. & Thrasher, A. J. Progress and prospects: gene therapy for inherited immunodeficiencies. *Gene Ther.* **16**, 1285–1291 (2009).
62. Stein, S. *et al.* Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. *Nat Med* **16**, 198–204 (2010).
63. Modlich, U. *et al.* Insertional transformation of hematopoietic cells by self-inactivating lentiviral and gammaretroviral vectors. *Mol Ther* **17**, 1919–1928 (2009).
64. Zychlinski, D. *et al.* Physiological promoters reduce the genotoxic risk of integrating gene vectors. *Mol Ther* **16**, 718–725 (2008).
65. Frecha, C., Szécsi, J., Cosset, F.-L. & Verhoeven, E. Strategies for targeting lentiviral vectors. *Curr. Gene Ther.* **8**, 449–460 (2008).
66. Cesana, D. *et al.* Uncovering and dissecting the genotoxicity of self-inactivating lentiviral vectors in vivo. *Mol. Ther.* **22**, 774–85 (2014).
67. Ginn, S. L., Alexander, I. E., Edelstein, M. L., Abedi, M. R. & Wixon, J. Gene therapy clinical trials worldwide to 2012 - an update. *Journal of Gene Medicine* **15**, 65–77 (2013).
68. Naldini, L. Gene therapy returns to centre stage. *Nature* **526**, 351–360 (2015).
69. Hacein-Bey Abina, S. *et al.* Outcomes following gene therapy in patients with severe Wiskott-Aldrich syndrome. *JAMA* **313**, 1550–63 (2015).
70. Hacein-Bey Abina, S. *et al.* A modified γ -retrovirus vector for X-linked severe combined immunodeficiency.

- N. Engl. J. Med.* **371**, 1407–17 (2014).
71. Hacein-Bey-Abina, S. *et al.* Efficacy of gene therapy for X-linked severe combined immunodeficiency. *N. Engl. J. Med.* **363**, 355–64 (2010).
 72. Hacein-bey-abina, S. *et al.* Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. - PubMed - NCBI. **118**, (2008).
 73. Aiuti, A. *et al.* Gene therapy for immunodeficiency due to adenosine deaminase deficiency. *N. Engl. J. Med.* **360**, 447–58 (2009).
 74. Malech, H. L. *et al.* Prolonged production of NADPH oxidase-corrected granulocytes after gene therapy of chronic granulomatous disease. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 12133–8 (1997).
 75. Ott, M. G. *et al.* Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. *Nat Med* **12**, 401–409 (2006).
 76. Aiuti, A. & Roncarolo, M. G. Ten years of gene therapy for primary immune deficiencies. *Hematol. Am Soc Hematol Educ Progr.* 682–689 (2009). doi:10.1182/asheducation-2009.1.682
 77. Candotti, F. Gene transfer into hematopoietic stem cells as treatment for primary immunodeficiency diseases. *Int. J. Hematol.* **99**, 383–392 (2014).
 78. Rawat, A., Bhattad, S. & Singh, S. Chronic Granulomatous Disease. *Indian J. Pediatr.* **83**, 345–53 (2016).
 79. Holland, S. M. Chronic Granulomatous Disease. *Clin. Rev. Allergy Immunol.* **38**, 3–10 (2009).
 80. Chiriaco, M., Salfa, I., Matteo, G. Di, Rossi, P. & Finocchi, A. Chronic Granulomatous Disease: clinical, molecular and therapeutic aspects. *Pediatr. Allergy Immunol.* 1–12 (2015). doi:10.1111/pai.12527
 81. Yu, L., Quinn, M., Cross, A. & Dinauer, M. Gp91 phox is the heme binding subunit of the superoxide-generating NADPH oxidase. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 7993–7998 (1998).
 82. Seger, R. A. Modern management of chronic granulomatous disease. *Br J Haematol* **140**, 255–266 (2008).
 83. Kaufmann, K. B. *et al.* Alpharetroviral Vector-mediated Gene Therapy for X-CGD: Functional Correction and Lack of Aberrant Splicing. *Mol. Ther.* **21**, 648–661 (2012).
 84. Kang, E. M. *et al.* Retrovirus gene therapy for X-linked chronic granulomatous disease can achieve stable long-term correction of oxidase activity in peripheral blood neutrophils. *Blood* **115**, 783–791 (2010).
 85. Grez, M. *et al.* Gene therapy of chronic granulomatous disease: the engraftment dilemma. *Mol. Ther.* **19**, 28–35 (2011).
 86. Schwäble, J., Schultze-Strasser, S., Stein, S. & Grez, M. [Gene therapy for septic granulomatosis: risks and opportunities. The road to successful treatment of congenital immunodeficiency]. *Pharm. Unserer Zeit* **40**, 264–271 (2011).
 87. Bianchi, M. *et al.* Restoration of NET formation by gene therapy in CGD controls aspergillosis. *Blood* **114**, 2619–2622 (2009).
 88. Kim JG, Ahn HS, Kang HJ, Kim S, Hong YT, J. C. Retroviral Gene Therapy for X-Linked Chronic Granulomatous Disease: Results from Phase I/II Trial. *Blood* **112**:2349, (2008).
 89. Stone, D., Niyonzima, N. & Jerome, K. R. Genome editing and the next generation of antiviral therapy. *Hum. Genet.* **135**, 1071–1082 (2016).
 90. Hoban, M. D. & Bauer, D. E. A genome editing primer for the hematologist. *Blood* **127**, blood-2016-01-678151 (2016).
 91. Johansen, S., Embley, T. M. & Willassen, N. P. A family of nuclear homing endonucleases. *Nucleic Acids Res.* **21**, 4405 (1993).
 92. Belfort, M. & Bonocora, R. P. Homing endonucleases: from genetic anomalies to programmable genomic clippers. *Methods Mol. Biol.* **1123**, 1–26 (2014).
 93. Hafez, M., Hausner, G. & Bonen, L. Homing endonucleases: DNA scissors on a mission. *Genome* **55**, 553–569 (2012).

94. Stoddard, B. L. Homing endonucleases: From microbial genetic invaders to reagents for targeted DNA modification. *Structure* **19**, 7–15 (2011).
95. Kim, Y. G., Cha, J. & Chandrasegaran, S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc Natl Acad Sci U S A* **93**, 1156–1160 (1996).
96. Gabriel, R. *et al.* An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat Biotechnol* **29**, 816–823 (2011).
97. Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S. & Gregory, P. D. Genome editing with engineered zinc finger nucleases. *Nat Rev Genet* **11**, 636–646 (2010).
98. Carroll, D. Genome engineering with zinc-finger nucleases. *Genetics* **188**, 773–782 (2011).
99. Boch, J. TALEs of genome targeting. *Nat. Publ. Gr.* **29**, 135–136 (2011).
100. Cermak, T. *et al.* Erratum: Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting (Nucleic Acids Research (2011) 39 (e82) DOI: 10.1093/nar/gkr218). *Nucleic Acids Res.* **39**, 7879 (2011).
101. Bogdanove, A. J., Schornack, S. & Lahaye, T. TAL effectors: finding plant genes for disease and defense. *Curr Opin Plant Biol* **13**, 394–401 (2010).
102. Mashimo, T. *et al.* Efficient gene targeting by TAL effector nucleases coinjected with exonucleases in zygotes. *Sci Rep* **3**, 1253 (2013).
103. Barrangou, R. *et al.* CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science (80-.)*. **315**, 1709–1712 (2007).
104. Ishino, Y., Shinagawa, H., Makino, K., Amemura, M. & Nakata, A. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.* **169**, 5429–33 (1987).
105. Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* **31**, 827–832 (2013).
106. Li, J. *et al.* Efficient inversions and duplications of mammalian regulatory DNA elements and gene clusters by CRISPR / Cas 9. **7**, 284–298 (2015).
107. Cong, L. *et al.* Multiplex Genome Engineering Using CRISPR/Cas Systems. **339**, 819–823 (2013).
108. Jinek, M. *et al.* RNA-programmed genome editing in human cells. *Elife* **2013**, 1–9 (2013).
109. Stella, S. & Montoya, G. The genome editing revolution: A CRISPR-Cas TALE off-target story. *Insid. Cell* n/a-n/a (2015). doi:10.1002/icl3.1038
110. Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262–1278 (2014).
111. Cox, D. B., Platt, R. J. & Zhang, F. Therapeutic genome editing: prospects and challenges. *Nat Med* **21**, 121–131 (2015).
112. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**, 2281–2308 (2013).
113. Gaj, T., Gersbach, C. A. & Barbas 3rd, C. F. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol* **31**, 397–405
114. Yin, H., Kauffman, K. J. & Anderson, D. G. Delivery technologies for genome editing. *Nat. Rev. Drug Discov.* **16**, 387–399 (2017).
115. Hart, T. *et al.* High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. *Cell* **163**, 1515–1526 (2015).
116. Tzelepis, K. *et al.* A CRISPR Dropout Screen Identifies Genetic Vulnerabilities and Therapeutic Targets in Acute Myeloid Leukemia. *Cell Rep.* **17**, 1193–1205 (2016).
117. Wang, T. *et al.* Gene Essentiality Profiling Reveals Gene Networks and Synthetic Lethal Interactions with Oncogenic Ras. *Cell* **168**, 890–903.e15 (2017).

118. Zhang, C. & Saunders, A. J. An emerging role for Ubiquilin 1 in regulating protein quality control system and in disease pathogenesis. *Discov. Med.* **8**, 18–22 (2009).
119. Young, C. S. *et al.* A Single CRISPR-Cas9 Deletion Strategy that Targets the Majority of DMD Patients Restores Dystrophin Function in hiPSC-Derived Muscle Cells. *Cell Stem Cell* **18**, 533–540 (2016).
120. Paquet, D. *et al.* Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* 1–18 (2016). doi:10.1038/nature17664
121. Xue, W. *et al.* CRISPR-mediated direct mutation of cancer genes in the mouse liver. *Nature* **514**, 380–384 (2014).
122. Flynn, R. *et al.* CRISPR-mediated genotypic and phenotypic correction of a chronic granulomatous disease mutation in human iPS cells. *Exp. Hematol.* **43**, 838–848.e3 (2015).
123. Wu, Y. *et al.* Correction of a genetic disease by CRISPR-Cas9-mediated gene editing in mouse spermatogonial stem cells. *Cell Res* **25**, 67–79 (2014).
124. Wu, Y. *et al.* Correction of a genetic disease in mouse via use of CRISPR-Cas9. *Cell Stem Cell* **13**, 659–662 (2013).
125. Li, H. *et al.* In vivo genome editing restores haemostasis in a mouse model of haemophilia. *Nature* **475**, 217–221 (2011).
126. Tebas, P. *et al.* Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N. Engl. J. Med.* **370**, 901–910 (2014).
127. Perez, E. E. *et al.* Editing Using Zinc-Finger Nucleases. **26**, 808–816 (2012).
128. Holt, N. *et al.* Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. *Nat Biotechnol* **28**, 839–847 (2010).
129. Ye, L. *et al.* Seamless modification of wild-type induced pluripotent stem cells to the natural CCR5 Δ 32 mutation confers resistance to HIV infection. *Proc. Natl. Acad. Sci.* **111**, 9591–9596 (2014).
130. Long, C. *et al.* Prevention of muscular dystrophy in mice by CRISPR/Cas9-mediated editing of germline DNA. *Science (80-.)*. **345**, 1184–1188 (2014).
131. Ousterout, D. G. *et al.* Reading frame correction by targeted genome editing restores dystrophin expression in cells from Duchenne muscular dystrophy patients. *Mol Ther* **21**, 1718–1726
132. Bloom, K., Ely, A., Mussolino, C., Cathomen, T. & Arbuthnot, P. Inactivation of Hepatitis B Virus Replication in Cultured Cells and In Vivo with Engineered Transcription Activator-Like Effector Nucleases. *Mol. Ther.* **21**, 1889–1897 (2013).
133. Lin, S.-R. *et al.* The CRISPR/Cas9 System Facilitates Clearance of the Intrahepatic HBV Templates In Vivo. *Mol. Ther. - Nucleic Acids* **3**, e186 (2014).
134. Genovese, P. *et al.* Targeted genome editing in human repopulating haematopoietic stem cells. *Nature* **510**, 235–240
135. Schwank, G. *et al.* Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell* **13**, 653–658 (2013).
136. Yin, H. *et al.* Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nat. Biotechnol.* **32**, 551–553 (2014).
137. Demaison, C. *et al.* High-level transduction and gene expression in hematopoietic repopulating cells using a human immunodeficiency [correction of imunodeficiency] virus type 1-based lentiviral vector containing an internal spleen focus forming virus promoter. *Hum. Gene Ther.* **13**, 803–813 (2002).
138. Subach, O. M. *et al.* Conversion of red fluorescent protein into a bright blue probe. *Chem. Biol.* **15**, 1116–124 (2008).
139. Engler, C., Kandzia, R. & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLoS One* **3**, (2008).

140. Abràmoff, M. D., Magalhães, P. J. & Ram, S. J. Image processing with ImageJ Part II. *Biophotonics Int.* **11**, 36–43 (2005).
141. Yamauchi, A. *et al.* Location of the epitope for 7D5, a monoclonal antibody raised against human flavocytochrome b558, to the extracellular peptide portion of primate gp91phox. *Microbiol. Immunol.* **45**, 249–57 (2001).
142. Brendel, C., Kaufmann, K. B., Krattenmacher, A., Pahujani, S. & Grez, M. Generation of X-CGD cells for vector evaluation from healthy donor CD34(+) HSCs by shRNA-mediated knock down of gp91(phox). *Mol Ther Methods Clin Dev* **1**, 14037 (2014).
143. Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* **11**, 783–784 (2014).
144. Zhen, L. *et al.* Gene targeting of X chromosome-linked chronic granulomatous disease locus in a human myeloid leukemia cell line and rescue by expression of recombinant gp91phox. *Proc Natl Acad Sci U S A* **90**, 9832–9836 (1993).
145. Urnov, F. D. *et al.* Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* **435**, 646–651 (2005).
146. Sebastiano, V. & Maeder, M. In situ genetic correction of the sickle cell anemia mutation in human induced pluripotent stem cells using engineered zinc finger nucleases. *Stem ...* **29**, 1717–1726 (2011).
147. Sun, N. & Zhao, H. Seamless correction of the sickle cell disease mutation of the HBB gene in human induced pluripotent stem cells using TALENs. *Biotechnol. Bioeng.* **111**, 1048–1053 (2014).
148. Xie, F. *et al.* Seamless gene correction of ??-thalassemia mutations in patient-specific iPSCs using CRISPR/Cas9 and piggyBac. *Genome Res.* **24**, 1526–1533 (2014).
149. Ma, N. *et al.* Transcription activator-like effector nuclease (TALEN)-mediated Gene correction in integration-free β -Thalassemia induced pluripotent stem cells. *J. Biol. Chem.* **288**, 34671–34679 (2013).
150. De Ravin, S. S. *et al.* CRISPR-Cas9 gene repair of hematopoietic stem cells from patients with X-linked chronic granulomatous disease. *Sci. Transl. Med.* **9**, (2017).
151. Choudhary, R., Baturin, D., Fosmire, S., Freed, B. & Porter, C. C. Knockdown of HPRT for selection of genetically modified human hematopoietic progenitor cells. *PLoS One* **8**, e59594 (2013).
152. Mao, Z., Bozzella, M., Seluanov, A. & Gorbunova, V. DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells. *Cell Cycle* **7**, 2902–6 (2008).
153. Newman, E. A. *et al.* Alternative NHEJ Pathway Components Are Therapeutic Targets in High-Risk Neuroblastoma. *Mol. Cancer Res.* **13**, 470–482 (2015).
154. Piñero, J., López-Baena, M., Ortiz, T. & Cortés, F. Apoptotic and necrotic cell death are both induced by electroporation in HL60 human promyeloid leukaemia cells. *Apoptosis* **2**, 330–336 (1997).
155. Mao, Z., Bozzella, M., Seluanov, A. & Gorbunova, V. DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells. *Cell Cycle* **7**, 2902–2906 (2008).
156. Beerman, I., Seita, J., Inlay, M. A., Weissman, I. L. & Rossi, D. J. Quiescent hematopoietic stem cells accumulate DNA damage during aging that is repaired upon entry into cell cycle. *Cell Stem Cell* **15**, 37–50 (2014).
157. Maggio, I. *et al.* Selection-free gene repair after adenoviral vector transduction of designer nucleases: Rescue of dystrophin synthesis in DMD muscle cell populations. *Nucleic Acids Res.* **44**, 1449–1470 (2015).
158. Jinnah, H. A., De Gregorio, L., Harris, J. C., Nyhan, W. L. & O'Neill, J. P. The spectrum of inherited mutations causing HPRT deficiency: 75 new cases and a review of 196 previously reported cases. *Mutat. Res. - Rev. Mutat. Res.* **463**, 309–326 (2000).
159. Valton, J. *et al.* Overcoming transcription activator-like effector (TALE) DNA binding domain sensitivity to cytosine methylation. *J. Biol. Chem.* **287**, 38427–38432 (2012).

160. Dreyer, A. *et al.* Biomaterials TALEN-mediated functional correction of X-linked chronic granulomatous disease in patient-derived induced pluripotent stem cells. *Biomaterials* **69**, 191–200 (2015).
161. Meaking, W. S., Edgerton, J., Wharton, C. W. & Meldrum, R. A. Electroporation-induced damage in mammalian cell DNA. *BBA - Gene Struct. Expr.* **1264**, 357–362 (1995).
162. Goldberg, A. & Rubinsky, B. The effect of electroporation type pulsed electric fields on DNA in aqueous solution. *Technol. Cancer Res. Treat.* **9**, 423–430 (2010).
163. Hoban, M. D. *et al.* Correction of the sickle cell disease mutation in human hematopoietic stem / progenitor cells. *Blood* **125**, 2597–2604 (2015).
164. Joglekar, A. V *et al.* Integrase-defective lentiviral vectors as a delivery platform for targeted modification of adenosine deaminase locus. *Mol Ther* **21**, 1705–17 (2013).
165. Nightingale, S. J. *et al.* Transient Gene Expression by Nonintegrating Lentiviral Vectors. *Mol. Ther.* **13**, 1121–1132 (2006).
166. Banasik, M. B. & McCray, P. B. Integrase-defective lentiviral vectors: progress and applications. *Gene Ther.* **17**, 150–157 (2010).
167. Wanisch, K. & Yáñez-Muñoz, R. J. Integration-deficient lentiviral vectors: a slow coming of age. *Mol. Ther.* **17**, 1316–1332 (2009).
168. Cornu, T. I. & Cathomen, T. Targeted Genome Modifications Using Integrase-deficient Lentiviral Vectors. *Mol. Ther.* **15**, 2107–2113 (2007).
169. Chew, W. L. *et al.* A multifunctional AAV–CRISPR–Cas9 and its host response. *Nat. Methods* **13**, 868–874 (2016).
170. Zarogoulidis, P. *et al.* Suicide Gene Therapy for Cancer – Current Strategies. *J Genet Syndr Gene Ther* (2013). doi:10.4172/2157-7412.1000139.Suicide
171. McCormick, F. Cancer Gene Therapy: Fringe or Cutting Edge? *Nat. Rev. Cancer* **1**, 130 (2001).
172. Saeger, J., Hytönen, V. P., Klotzsch, E. & Vogel, V. GFP's Mechanical Intermediate States. *PLoS One* **7**, (2012).
173. Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E. S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science (80-.)*. **343**, 80–84
174. Fu, Y. *et al.* High frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol* **31**, 822–826 (2014).
175. Dinuer, M. C., Gifford, M. A., Pech, N., Li, L. L. & Emshwiller, P. Variable correction of host defense following gene transfer and bone marrow transplantation in murine X-linked chronic granulomatous disease. *Blood* **97**, 3738–3745 (2001).
176. Kuhns, D. B. *et al.* Residual NADPH oxidase and survival in chronic granulomatous disease. *N. Engl. J. Med.* **363**, 2600–10 (2010).
177. Vouillot, L., Thelie, A. & Pollet, N. Comparison of T7E1 and surveyor mismatch cleavage assays to detect mutations triggered by engineered nucleases. *G3* **5**, 407–415 (2015).
178. Roos, D. *et al.* Hematologically important mutations: X-linked chronic granulomatous disease (third update). *Blood Cells. Mol. Dis.* **45**, 246–65 (2010).
179. O'Neill, S., Brault, J., Stasia, M. J. & Knaus, U. G. Genetic disorders coupled to ROS deficiency. *Redox Biol.* **6**, 135–156 (2015).
180. Cross, A. R., Rae, J. & Curnutte, J. T. Cytochrome b-245 of the neutrophil superoxide-generating system contains two nonidentical hemes. Potentiometric studies of a mutant form of gp91phox. *J Biol Chem* **270**, 17075–17077 (1995).
181. Cross, A. R., Heyworth, P. G., Rae, J. & Curnutte, J. T. A variant X-linked chronic granulomatous disease patient (X91+) with partially functional cytochrome b. *J. Biol. Chem.* **270**, 8194–200 (1995).

182. Stroncek, D. & McCullough, J. Safeguarding the long-term health of hematopoietic stem cell donors: a continuous and evolving process to maintain donor safety and trust. *Expert Rev. Hematol.* **5**, 1–3 (2012).
183. CYBBbase database. <http://structure.bmc.lu.se/idbase/CYBBbase/browser.php?content=browser> (2017).
184. Puck, J. M. *et al.* Mutation analysis of IL2RG in human X-linked severe combined immunodeficiency. *Blood* **89**, 1968–77 (1997).
185. IL2RG database. https://research.cchmc.org/LOVD2/home.php?select_db=IL2RG (2017).
186. WASP database. <http://pidj.rcai.riken.jp/wasbase/> (2017).
187. Jin, Y. *et al.* Mutations of the Wiskott-Aldrich Syndrome Protein (WASP): Hotspots, effect on transcription, and translation and phenotype/genotype correlation. *Blood* **104**, 4010–4019 (2004).
188. ADAbase Database. <http://structure.bmc.lu.se/idbase/ADAbase/index.php?content=index/IDbases> (2017).
189. Giardine, B. *et al.* Updates of the HbVar database of human hemoglobin variants and thalassemia mutations. *Nucleic Acids Res.* **42**, 1–10 (2014).
190. HBB Database. <https://lovd.bx.psu.edu/home.php> (2017).
191. Wefers, B. *et al.* Direct production of mouse disease models by embryo microinjection of TALENs and oligodeoxynucleotides. *Proc. Natl. Acad. Sci.* **110**, 3782–3787 (2013).
192. DeWitt, M. A., Corn, J. E. & Carroll, D. Genome editing via delivery of Cas9 ribonucleoprotein. *Methods* **3**–9 (2017). doi:10.1016/j.ymeth.2017.04.003
193. Hendel, A. *et al.* Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat. Biotechnol.* **33**, (2015).

Acknowledgments

Concluding my dissertation, I would like to thank all the people who contributed the success of this work during the last years.

First, I would like to express my special gratitude and thanks to my supervisor Prof. Dr. Harald von Melchner, he has been a highly appreciated mentor for me. I would like to thank him for encouraging my research and for helping me to grow as a research scientist.

In particular I would like to thank Prof. Dr. Beatrix Süß and Prof. Dr. M. Cristina Cardoso for being my supervisors on the TU Darmstadt and for thier kind advising, supporting and discussing my doctoral thesis.

I am deeply grateful to PD. Dr. Frank Schnütgen for his mentoring during my work in the laboratory, for his directional instructions, advices and valuable discussion about my doctoral dissertation.

It's my great pleasure to express my appreciation to Prof. Dr. Hubert Serve, head of the Department of Hematology and Oncology Goethe-University Frankfurt, for giving me the opportunity to conduct my doctoral study in the field of Hematology.

Furthermore, I am greatly indebted to express my deepest thanks to a number of individuals who deserve special acknowledgement: Ms. Anna-Maria Scheder for the great support in all matters, without her many things would have been heavier; Dr. Ana Tomasovic for her advices, generous friendship, psychological support and excellent proof reading of the manuscript; Dr. Stefan Stein for his continuous encouragement; Many others at our department who have also been generous with their time, support and encouragement at various stages during my PhD including Dr. Frank Wempe, Dr. Nina Kurrle, Dr. Olesya Vakhrusheva, Sarah Weber, Florian Gatzke, Christine Schönfeld, Jenny Bleeck, Heike Nürnberger, Johanna Kreitz, Melanie Forster, Roy Ehling, The Duy Nguyen, Nathalie Guillen, Sebastian Koschade and other colleagues for the support and the nice atmosphere.

I would like to thank my friends Sabine Arens, Ali Bildik and Meded Bildik. They have been there forever helping me with their wisdoms.

Finally, heartfelt thanks to my parents Mustafa and Elif Sürün, my sisters Hamide, Fatma, Emine and my brother Ali for their continuous love and encouragement throughout the years.

Curriculum vitae

Duran Sürün

Address: Lerchenstraße 6
63150 Heusenstamm

Phone number: +49178-2314333

E-mail: d.sueruen@mail.de

Date of birth: January 1st 1983

Nationality: German



WORK EXPERIENCE & EDUCATION

Mai 2012 – present	PhD Thesis at Goethe-University in the department of Hematology and Oncology, Frankfurt am Main In the Laboratory of Prof. Dr. Harald von Melchner. Dissertation title: "High Efficiency Gene Correction in Hematopoietic Cells by Donor Template-free CRISPR/Cas9 Genome Editing"
Oct 2011 – Mai 2012	Scientific assistant at Georg-Speyer-Haus In the Laboratory of Dr. Manuel Grez, Dr. Manuel Grez
Oct 2010 – Oct 2011	Master thesis in Georg-Speyer-Haus In the Laboratory of Dr. Manuel Grez Grade 1.1 Master title: „Sleeping Beauty – Vektoren für die genetische Modifikation hämatopoetischer Stammzellen“
Oct 2009 – Oct 2011	Master of science at University Gelsenkirchen in Molecular Biology Grade 1.4
Apr 2009 – Aug 2009	Bachelor thesis at Heinrich-Pette Institut In the Laboratory of Prof. Dr. Wolfgang Deppert Grade 1.0
Oct 2006 – Sept 2009	Bachelor of science in Molecular Biology, University Gelsenkirchen Grade 2.3
Sept 2003 – Jun 2006	Training as Chemical-technical assistant at Hans-Böckler Berufskolleg, Marl Grade 2.5
Sept 2003 – Jun 2006	High school graduation at Hans-Böckler Berufskolleg, Marl Grade 2.5

TEACHING EXPERIENCE

Apr 2008 – Dec 2009	Student Assistant in University Gelsenkirchen in the department of biochemistry, microbiology and laboratory practice
Feb 2015	Students seminar "The renewable man?" at the EKHN Stiftung, Frankfurt am Main, Germany

Publications and Abstracts

CONFERENCES AND AWARDS

Aug 2017	Oral presentation at the 46 th Annual scientific International Society for Experimental Hematology (ISEH) meeting, Frankfurt, Germany
Dec 2016	Poster presentation at the 58 th Annual American Society for Hematology (ASH) meeting, San Diego, USA
Nov 2016	Poster presentation at the Nature Herrenhausen Symposium Genome Editing for Gene and Cell Therapy meeting, Hanover, Germany
Oct 2015	Poster prize at the Summer School of the LOEWE Centre for Cell and Gene Therapy, Langen, Germany
Form 2012 – 2016	Poster presentation at the Annual meeting of the German Society for Gene Therapy (DG-GT)

PUBLICATION

Sürün D, Schwäble J, Tomasovic A, Roy E, Stein S, Kurrle A, Kühn A, von Melchner H and Schnütgen F

High Efficiency Gene Correction in Hematopoietic Cells by Donor-Template-free CRISPR/Cas9 Genome Editing

Molecular Therapy - Nucleic Acids, in Revision

Cencioni C, Spallotta F, Savoia M, Kuenne C, Guenther S, Re A, Wingert S, Rehage M, Sürün D, Siragusa M, Smith JG, Schnütgen F, von Melchner F, Rieger M, Martelli F, Riccio A, Fleming I, Braun T, Zeiher AM, Farsetti A and Gaetano C

A Zeb1 Hdac2 eNOS feedback circuitry identifies early cardiovascular precursors in naïve mouse embryonic stem cells

Nature Communications, in Revision

Spallotta F, Cencioni C, Atlante S, Garella D, Cocco M, Mori M, Mastrocola R, Kuenne C, Guenther S, Nanni S, Azzimato V, Zukunft S, Kornberger A, Sürün D, Schnütgen F, von Melchner H, Di Stilo A, Aragno M, De Blasio M, Ritchie RH, Zaccagnini G, Martelli F, Farsetti A, Fleming I, Braun T, Beiras-Fernandez A, Botta B, Collino M, Bertinaria M, Zeiher AM, Gaetano C

Stable oxidative cytosine modifications accumulate in cardiac mesenchymal cells from Type2 diabetes patients: rescue by alpha-ketoglutarate and TET-TDG functional reactivation

Circulation Research, submitted

Nguyen TD, Shaid S, Vakhrusheva O, Koschade S, Zhang J, Oellerich T, Sürün D, Derlet A, Osiewacz H, Behrends C, Dikic I, and Brandts C

The selective autophagy receptor p62 is required for mitophagy and promotes leukemia development

The Journal of Experimental Medicine, submitted

Ball AK, Beilstein K, Wittmann S, Sürün D, Saul MJ, Schnütgen F, Flamand N, Capelo R, Kahnt AS, Frey H, Schaefer L, Marschalek R, Häfner AK, Steinhilber D

Characterization and cellular localization of human 5-lipoxygenase and its protein isoforms 5-LO Δ 13, 5-LO Δ 4 and 5-LOp12

Biochim Biophys Acta.2017; 1862(5):561-571

Tomasovic A, Kurrle N, Wempe F, De-Zolt S, Scheibe S, Koli K, Serchinger M, Schnütgen F, Sürün D, Sterner-Kock A, Weissmann N, von Melchner H

Ltbp4 regulates Pdgfr β expression via TGF β -dependent modulation of Nrf2 transcription factor function

Matrix Biol. 2017; 59:109-120.

Mahshid Y, Markoutsa S, Dincbas-Renqvist V, Sürün D, Christensson B, Sander B, Björkholm M, Sorg BL, Rådmark O, Claesson HE

Phosphorylation of serine 523 on 5-lipoxygenase in human B lymphocytes

Prostaglandins Leukot Essent Fatty Acids 2015; 100:33-40

Tomasovic A, Kurrle N, Sürün D, Heidler J, Husnjak K, Poser I, Schnütgen F, Scheibe S, Seimetz M, Jaksch P, Hyman A, Weissmann N, von Melchner H

Sestrin 2 protein regulates platelet-derived growth factor receptor β (Pdgfr β) expression by modulating proteasomal and Nrf2 transcription factor functions

J Biol Chem. 2015; 10;290(15):9738-52

Markoutsa S, Sürün D, Karas M, Hofmann B, Steinhilber D, Sorg BL

Analysis of 5-lipoxygenase phosphorylation on molecular level by MALDI-MS

FEBS J. 2014; 281(8):1931-47

Ehrenwörtliche Erklärung:

Ich erkläre hiermit ehrenwörtlich, dass ich die vorliegende Arbeit entsprechend den Regeln guter wissenschaftlicher Praxis selbstständig und ohne unzulässige Hilfe Dritter angefertigt habe.

Sämtliche aus fremden Quellen direkt oder indirekt übernommenen Gedanken sowie sämtliche von Anderen direkt oder indirekt übernommenen Daten, Techniken und Materialien sind als solche kenntlich gemacht. Die Arbeit wurde bisher bei keiner anderen Hochschule zu Prüfungszwecken eingereicht.

Darmstadt, den 01.11.2017

.....

Duran Sürün